



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 180133

TO: Ruixiang Li
Location: 4c75 / 4c70
Tuesday, February 28, 2006
Art Unit: 1646
Phone: 571-272-0875
Serial Number: 10 / 618570

From: Jan Delaval
Location: Biotech-Chem Library
Remsen 1a51
Phone: 571-272-2504

jan.delaval@uspto.gov

Search Notes

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

STIC-Biotech/ChemLib

180133

my

From: Li, Ruixiang
Sent: Tuesday, February 21, 2006 4:59 PM
To: STIC-Biotech/ChemLib
Subject: Sequence search of Application No.10/618570

Please do a standard search on:

SEQ ID NOS: 1 and 2 against nucleic acid databases (excluding pending databases).

Thank you very much!

Ruixiang Li
GAU 1646
REM 4D75
Mail Box 4C70
(571) 272-0875

STIC
2006 22 2006

Searcher: Jan
Searcher Phone: 22504
Date Searcher Picked up: 2/22/06
Date completed: 2/23/06
Searcher Prep Time: 10
Online Time: 15

Type of Search
NA# ✓ AA#
S/L: Oligomer:
Encode/Transl: ✓
Structure #: Text:
Inventor: Litigation:

Vendors and cost where applicable
STN:
DIALOG:
QUESTEL/ORBIT:
LEXIS/NEXIS:
SEQUENCE SYSTEM: ✓
WWW/Internet:
Other (Specify):

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (C) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2006, 17:05:18 ; Search time 47 Seconds
(without alignments)
703.623 Million cell updates/sec

Title: US-10-618-570-2

Perfect score: 2059

Sequence: 1 MAQWDDPPQDQEDTSCSTES.....DKATRVGINFTLRQKE 400

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptcdat1/1/aa/5 COMB.pep:
2: /cgn2_6/ptcdat1/1/aa/5 COMB.pep:
3: /cgn2_6/ptcdat1/1/aa/5 COMB.pep:
4: /cgn2_6/ptcdat1/1/aa/5 COMB.pep:
5: /cgn2_6/ptcdat1/1/aa/5 COMB.pep:
6: /cgn2_6/ptcdat1/1/aa/5 COMB.pep:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1381	67.1	349	6	5510466-2
2	1381	67.1	453	6	5510466-4
3	1047.5	50.9	451	1	US-08-154-365-2
4	1042.5	50.6	358	1	US-08-453-117-4
5	1042.5	50.6	358	1	US-08-948-222-4
6	1042.5	50.6	358	1	US-08-973-145-4
7	1042.5	50.6	358	4	PCT-US96-08081-4
8	1042.5	50.6	362	2	US-09-949-016-9258
9	1042.5	50.6	451	1	US-08-453-117-2
10	1042.5	50.6	451	1	US-08-948-222-2
11	1042.5	50.6	451	1	US-08-973-145-2
12	1042.5	50.6	451	2	US-09-276-400-10
13	1042.5	50.6	451	2	US-09-448-076-10
14	1042.5	50.6	451	2	US-09-702-572-10
15	1042.5	50.6	451	4	PCT-US96-08081-2
16	678	32.9	152	4	US-08-831-399-4
17	678	32.9	152	2	US-09-368-862-4
18	678	32.9	152	2	US-09-368-772-4
19	138.5	6.7	159	2	US-08-628-540-1
20	138.5	6.7	159	2	US-08-948-097-17
21	138.5	6.7	159	2	US-08-941-100-1
22	138.5	6.7	159	2	US-08-941-100-5
23	138.5	6.7	159	2	US-08-381-430-2
24	138.5	6.7	159	2	US-08-831-399-2
25	138.5	6.7	183	2	US-09-368-862-2
26	138.5	6.7	183	2	US-09-368-772-2
27	138.5	6.7	183	4	PCT-US93-05240-14

28	138.5	6.7	183	6	5168049-5	Patent No. 5168049
29	138.5	6.7	415	1	US-08-491-988-7	Sequence 7, Appl
30	138.5	6.7	421	2	US-09-338-270B-1	Sequence 1, Appl
31	138.5	6.7	435	1	US-08-491-988-5	Sequence 5, Appl
32	138.5	6.6	128	1	US-08-411-833-2	Sequence 2, Appl
33	138.5	6.6	128	1	US-08-434-718-2	Sequence 2, Appl
34	138.5	6.6	128	2	US-08-831-399-16	Sequence 16, Appl
35	138.5	6.6	128	2	US-09-368-862-16	Sequence 16, Appl
36	138.5	6.6	128	2	US-09-368-772-16	Sequence 16, Appl
37	138.5	6.6	402	1	US-08-491-988-9	Sequence 9, Appl
38	138.5	6.6	10182	2	US-09-134-001C-3159	Sequence 3159, Ap
39	133	6.5	185	2	US-09-382-276-2	Sequence 2, Appl
40	132.5	6.4	118	2	US-07-780-717C-7	Sequence 7, Appl
41	131	6.4	186	2	US-09-382-276-3	Sequence 3, Appl
42	130.5	6.3	183	2	US-09-382-276-1	Sequence 1, Appl
43	130.5	6.3	183	2	US-09-485-867-1	Sequence 1, Appl
44	129.5	6.3	1307	2	US-09-949-016-7561	Sequence 7561, Ap
45	123	6.0	2663	2	US-09-538-092-1252	Sequence 1252, Ap

ALIGNMENTS

RESULT 1
5510466-2
; Patent No. 5510466
; APPLICANT: KREIGER, MONTY; KODAMA, TATSUHIKO
; TITLE OF INVENTION: SCAVENGER RECEPTOR PROTEIN AND ANTIBODY
; THERE TO
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,400
; FILING DATE: 16-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 997,113
; FILING DATE: 24-DEC-1992
; APPLICATION NUMBER: 391,486
; FILING DATE: 09-AUG-1989
; APPLICATION NUMBER: 272,002
; FILING DATE: 15-NOV-1988
; SEQ ID NO.2:
; LENGTH: 349
5510466-2

Query Match	67.1%	Score	1381	DB	6	Length	349
Best Local Similarity	100.0%	Pred. No.	4.3e-112				
Matches	272	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	MAQWDDPPQDQEDTSCSTESVKF	DARSVTALLPPHPKNGPTLQERMKSVK	TALITLYLIV	60		
Db	1	MAQWDDPPQDQEDTSCSTESVKF	DARSVTALLPPHPKNGPTLQERMKSVK	TALITLYLIV	60		
Qy	61	FVVLVPIIGIVAAQLKWKETKNC	TGVSVNADISPSPEGKNGSDEMR	PFRAVMERMSNM	120		
Db	61	FVVLVPIIGIVAAQLKWKETKNC	TGVSVNADISPSPEGKNGSDEMR	PFRAVMERMSNM	120		
Qy	121	ESRQYLSNDEANLLDAKNFQNF	SITTDQRFNDVLPQLNSLS	SIQEHENIIGDISKSLV	180		
Db	121	ESRQYLSNDEANLLDAKNFQNF	SITTDQRFNDVLPQLNSLS	SIQEHENIIGDISKSLV	180		
Qy	181	GLNTTVLDLQPSITELNGRVQEN	APKQSEMKLBERIYNASAEIK	SDKQVLEOEIK	240		
Db	181	GLNTTVLDLQPSITELNGRVQEN	APKQSEMKLBERIYNASAEIK	SDKQVLEOEIK	240		
Qy	241	GEMKLLNNTNDRKDKWEHSQ	TLRKNITLLQ	272			
Db	241	GEMKLLNNTNDRKDKWEHSQ	TLRKNITLLQ	272			

RESULT 2
5510466-4
; Patent No. 5510466
; APPLICANT: KREIGER, MONTY; KODAMA, TATSUHIKO

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2006, 08:21:52 ; Search time 18619 Seconds
(without alignments)
13009.116 Million cell updates/sec

Title: US-10-618-570-1
Perfect score: 5177
Sequence: 1 ttgaagaccaccaccgta.....agcgggggtcttcattgg 5177

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_hic:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_est7:*

9: gb_gsa1:*

10: gb_gsa2:*

11: gb_gsa3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	829	16.0	935	7	CK284786
2	812.4	15.7	856	7	CK287297
3	812.4	15.7	910	7	CK287930
4	812.4	15.7	933	7	CK291799
5	812.4	15.7	936	7	CK256977
6	812.4	15.7	947	7	CK298208
7	812.4	15.7	954	7	CK283361
8	803	15.5	804	7	CK291519
9	794	15.3	811	7	CK288185
10	793	15.3	793	8	DN548862
11	737.4	14.2	767	8	DR696294
12	707.2	13.7	793	8	CF409555
13	634.2	12.3	837	10	C2169992
14	605.4	11.7	878	7	CK288911
15	589.4	11.4	591	8	DR696887
16	588.8	11.4	789	9	AQ361914
17	582.8	11.3	586	7	CK801096
18	576.4	11.1	620	9	AQ398387
19	566.6	10.9	1033	10	AY413208
20	551.2	10.6	1027	10	AY413209
21	542.4	10.5	549	7	CK801630
22	541.2	10.5	804	8	DR104979

23	538.2	10.4	978	7	CN643776
24	535	10.3	580	8	DR108385
25	531.4	10.3	787	1	AM037916
26	530	10.2	759	6	CB172376
27	523.8	10.1	637	9	AQ447230
28	522.2	10.1	592	9	AQ49162
29	514.8	9.9	561	9	AQ447874
30	510.6	9.9	843	5	EX498563
31	508.4	9.8	562	6	CB453499
32	508	9.8	509	9	AQ447775
33	502.4	9.7	652	6	CB453082
34	499	9.6	499	9	AQ398880
35	496.4	9.6	498	8	DR696295
36	494.2	9.6	498	9	AQ398730
37	494.2	9.5	857	9	AQ875013
38	489.4	9.4	491	8	DR696578
39	485.2	9.4	551	1	AJ677598
40	483.6	9.3	850	9	AQ875052
41	483	9.3	707	8	DR108377
42	481.8	9.3	798	9	AQ876139
43	479.6	9.3	740	10	AG002307
44	475.2	9.2	861	9	AQ875040
45	474.2	9.2	710	9	BH972732

ALIGNMENTS

RESULT 1
CK284786
LOCUS
DEFINITION
EST747508 Nicotiana benthamiana mixed tissue cDNA library,
normalized, full-length Nicotiana benthamiana cDNA clone NBMAQ41 5',
end, mRNA sequence.
ACCESSION
CK284786
VERSION
CK284786.1
KEYWORDS
EST.
SOURCE
Nicotiana benthamiana
ORGANISM
Nicotiana benthamiana
REFERENCE
1 (bases 1 to 935)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,
Staskawicz, B., Jin, H. and Baker, B.
Generation of EST sequences from Nicotiana benthamiana
Other ESTs: EST747509
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/
Seq primer: ATT TAG GTG ACA CTA TAG.
Location/Qualifiers
1 .935
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clones="NBMAQ41"
/tissue_type="abiotic and biotic stress-treated leaves,
callus tissue and root tissue"
/lab_hosts="DH10B-Tona"
/clone_lib="Nicotiana benthamiana mixed tissue cDNA
library, normalized, full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Nicotiana benthamiana
tissues that include callus, roots from liquid culture
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
challenged leaves (Pseudomonas syringae pv tomato 12 hr;
Xanthomonas campestris pv campestris 12 hr, 18hr;

Pseudomonas syringae pv *phaseolicola* 18hr, and *Xanthomonas campestris* pv *vesicatoria* 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."

```
ORIGIN
Query Match      16.0%; Score 829; DB 7; Length 935;
Best Local Similarity 99.4%; Pred. No. 1.5e-215;
Matches 832; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3706 ACAGATGAGGATCGTTTCGATGATGAACAAGATGATGACGACGAGGTTCTCCGGCC 3765
DB 11 ACCGATCTGATCGTTTCGATGATGAACAAGATGATGACGACGAGGTTCTCCGGCC 70

QY 3766 GCTTGGGTGGAGAGCTATTCCGCTATCATCTGGGCAACAAGACATCGGCTGCTGAT 3825
DB 71 GCTTGGGTGGAGAGCTATTCCGCTATCATCTGGGCAACAAGACATCGGCTGCTGAT 130

QY 3826 CGCGCGGTGTTCCGGCTGTGACGCGAGGGCGCGCGTTCCTTTTGTCAAGACGACCTG 3885
DB 131 CGCGCGGTGTTCCGGCTGTGACGCGAGGGCGCGCGTTCCTTTTGTCAAGACGACCTG 190

QY 3886 TCCGCTGCCCTGAATGAATCTGACGACGAGGCGCGGCTATCGTGCTGGCCACGACG 3945
DB 191 TCCGCTGCCCTGAATGAATCTGACGACGAGGCGCGGCTATCGTGCTGGCCACGAGC 250

QY 3946 GCGGTTCTTGGCGAGCTGTGCTGACGTTGTCACTGAAGCGGAGAGGACTGCTGCTA 4005
DB 251 GCGGTTCTTGGCGAGCTGTGCTGACGTTGTCACTGAAGCGGAGAGGACTGCTGCTA 310

QY 4006 TTGGSCGAAGTGC CGGGCGAGGATCTCTGTCTCATCTCACTTGTCTCTCGCGGAGAAAGTA 4065
DB 311 TTGGSCGAAGTGC CGGGCGAGGATCTCTGTCTCATCTCACTTGTCTCTCGCGGAGAAAGTA 370

QY 4066 TCCATCATGCTGATGCAATCGCGGCTGTGATAGCTTGTATCGGGTACCTGCGCCATTC 4125
DB 371 TCCATCATGCTGATGCAATCGCGGCTGTGATAGCTTGTATCGGGTACCTGCGCCATTC 430

QY 4126 GACCAACGAGGAAACATCGATCGAGCGAGCACTACTCGGATGGAAGCGGCTTGTGTC 4185
DB 431 GACCAACGAGGAAACATCGATCGAGCGAGCACTACTCGGATGGAAGCGGCTTGTGTC 490

QY 4186 GATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCGCACTGTCGCCAGG 4245
DB 491 GATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCGCACTGTCGCCAGG 550

QY 4246 CTCAGGCGCGCATGCCCGACGCGGAGGATCTCGTGTGACCCATGCGGATGCTGCTTG 4305
DB 551 CTCAGGCGCGCATGCCCGACGCGGAGGATCTCGTGTGACCCATGCGGATGCTGCTTG 610

QY 4306 CGGAATATCATGTTGGAATAATGGCGCTTTTCTGATTCATCGACTGTGGCGGCTGGGT 4365
DB 611 CGGAATATCATGTTGGAATAATGGCGCTTTTCTGATTCATCGACTGTGGCGGCTGGGT 670

QY 4366 GTGGCGGACCGCTATCAGACATACGCTTGGCTACCCGTGATATGCTGAAGAGCTTGGC 4425
DB 671 GTGGCGGACCGCTATCAGACATACGCTTGGCTACCCGTGATATGCTGAAGAGCTTGGC 730

QY 4426 GCGGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTCGACGCGC 4485
DB 731 GCGGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTCGACGCGC 790

QY 4486 ATGCGCTTCTATCGCTTCTTGAAGAGTTCTTCTGAGCGGAGCTCTGGGTTTGATA 4542
DB 791 ATGCGCTTCTATCGCTTCTTGAAGAGTTCTTCTGAGCGGAGCTCTGGGTTTGATA 847
```

RESULT 2
CK287297 856 bp mRNA linear EST 02-AUG-2004
LOCUS
DEFINITION
EST750019 *Nicotiana benthamiana* mixed tissue cDNA library,
normalized, full-length *Nicotiana benthamiana* cDNA clone NBM815 5'
end, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CK287297
CK287297.1 GI:39863696
EST.
Nicotiana benthamiana
Nicotiana benthamiana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; *Nicotiana*.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 856)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,
Staskawicz, B., Jin, H. and Baker, B.
Generation of EST sequences from *Nicotiana benthamiana*
Unpublished (2003)
Other ESTs: EST750020

Contact: Robin Buell

The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org

Clones can be requested from the University of Arizona Genomics
Institute via <http://genome.arizona.edu/orders/>.
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

source

1..856
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clone="NBM815"
/tissue_type="abiotic and biotic stress-treated leaves,
callus tissue and root tissue"
/lab_host="DH10B-TonA"
/clone_lib="Nicotiana benthamiana mixed tissue cDNA
library, normalized, full-length"
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from *Nicotiana benthamiana*
tissues that include callus, roots from liquid culture
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
challenged leaves (*Pseudomonas syringae* pv *tomato* 12 hr;
Xanthomonas campestris pv *campestris* 12 hr, 18hr;
Pseudomonas syringae pv *phaseolicola* 18hr, and *Xanthomonas*
campestris pv *vesicatoria* 18hr). RNA was isolated from
these tissues and pooled in approximately equal molar
amounts."

ORIGIN

Query Match 15.7%; Score 812.4; DB 7; Length 856;
Best Local Similarity 99.9%; Pred. No. 5.2e-211;
Matches 813; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 3729 GATTGACAAAGATGGATTGACACGAGGTTCTCGGCGCTTGGTGAGAGGCTATTTCGG 3788
DB 14 GATTGACAAAGATGGATTGACACGAGGTTCTCGGCGCTTGGTGAGAGGCTATTTCGG 73

QY 3789 CTATGACTGGGCAACAAGACAAATCGGCTGCTCTGATGCCCGCTGTTCGGCTGTGAGC 3848
DB 74 CTATGACTGGGCAACAAGACAAATCGGCTGCTCTGATGCCCGCTGTTCGGCTGTGAGC 133

QY 3849 GCGAGGGCGCGCGGTTCTTTTGTCAAGACGAGCTGTCCGCTGCCCTGATGAATGCA 3908
DB 134 GCGAGGGCGCGCGGTTCTTTTGTCAAGACGAGCTGTCCGCTGCCCTGATGAATGCA 193

QY 3909 GGACGAGGCGCGGCTATCGTGCTGGCCACGACGCGGGGTTCTTTCGCGAGCTGTGCT 3968
DB 194 GGACGAGGCGCGGCTATCGTGCTGGCCACGACGCGGGGTTCTTTCGCGAGCTGTGCT 253

QY 3969 CGAGCTTGTCACTGAACGCGGAGGAGCTGGCTGCTATTGGGCGAAGTTCGCGGCGCAGGA 4028
DB 254 CGAGCTTGTCACTGAACGCGGAGGAGCTGGCTGCTATTGGGCGAAGTTCGCGGCGCAGGA 313

QY 4029 TCTCCTCTCATCTCACCTTCTCTGCGGAGAAAGTATCCATCATATGCTGATGCAATGCG 4088
DB 314 TCTCCTCTCATCTCACCTTCTCTGCGGAGAAAGTATCCATCATATGCTGATGCAATGCG 373
```

4089 CGGGCTGATACGCTTGTATCGGCTACCTGCCATTGACCAACGGAACATCGCAT 4148
|||||
Db |||||
374 GCGGCTGATACGCTTGTATCGGCTACCTGCCATTGACCAACGGAACATCGCAT 433
|||||
Qy |||||
4149 CGAGCGACGATCTACTCGGATGGAAGCGGCTTGTGCGATCAGGATGATCTGGACGAAGA 4208
|||||
Db |||||
434 CGAGCGACGATCTACTCGGATGGAAGCGGCTTGTGCGATCAGGATGATCTGGACGAAGA 493
|||||
Qy |||||
4209 GCATCAGGGGCTCGCGCAGCGCAACTGTTGCGCAGGCTCAAGCGCGCATGCGCGACGG 4268
|||||
Db |||||
494 GCATCAGGGGCTCGCGCAGCGCAACTGTTGCGCAGGCTCAAGCGCGCATGCGCGACGG 553
|||||
Qy |||||
4269 CGAGGATCTCTGTCGACCCATGCGGATGCGCTTGGCGGATATCATGTTGGAATGG 4328
|||||
Db |||||
554 CGAGGATCTCTGTCGACCCATGCGGATGCGCTTGGCGGATATCATGTTGGAATGG 613
|||||
Qy |||||
4329 CCGCTTTTCTCGATTCATCGACTGTGGCGCGCTGGGTGCGGACCGCTATCAGGACAT 4388
|||||
Db |||||
614 CCGCTTTTCTCGATTCATCGACTGTGGCGCGCTGGGTGCGGACCGCTATCAGGACAT 673
|||||
Qy |||||
4389 AGCGTTGGCTACCGCTGATATGCTGAAGAGCTTGGCGGCAATGGGCTGACCGCTTCT 4448
|||||
Db |||||
674 AGCGTTGGCTACCGCTGATATGCTGAAGAGCTTGGCGGCAATGGGCTGACCGCTTCT 733
|||||
Qy |||||
4449 CGTGTTCATCGGATCGCGCTCCGATTCGAGGCGATCCCTTCTATCGCCTTCTGA 4508
|||||
Db |||||
734 CGTGTTCATCGGATCGCGCTCCGATTCGAGGCGATCCCTTCTATCGCCTTCTGA 793
|||||
Qy |||||
4509 CGAGTTCTTCGAGCGGACTCTGGGTTTCGATA 4542
|||||
Db |||||
794 CGAGTTCTTCGAGCGGACTCTGGGTTTCGATA 827
|||||

RESULT 3

CK287930
LOCUS EST750652 Nicotiana benthamiana mixed tissue cDNA library,
normalized, full-length Nicotiana benthamiana cDNA clone NBMBCT5 5'
end, mRNA sequence.

ACCESSION CK287930.1 GI:39864940

VERSION EST.

KEYWORDS

SOURCE

ORGANISM

Nicotiana benthamiana
Nicotiana benthamiana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.

REFERENCE 1 (bases 1 to 910)

Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,
Staskiewicz, B., Jin, H. and Baker, B.

Generation of EST sequences from Nicotiana benthamiana

Unpublished (2003)

Other ESTs: EST750653

Contact: Robin Suell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from the University of Arizona Genomics

Institute via <http://genome.arizona.edu/orders/>

Seq primer: ATT TAG GTG ACA CTA TAG.

Location/Qualifiers

1. 910

source

/organism="Nicotiana benthamiana"

/mol_type="mRNA"

/db_xref="taxon:4100"

/clone="NBMBCT5"

/tissue_type="abiotic and biotic stress-treated leaves,"

callus tissue and root tissue"

/lab_host="DH10B-TonA"

/clone_lib="Nicotiana benthamiana mixed tissue cDNA

library, normalized, full-length"

/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;

supplier: RNA was isolated from Nicotiana benthamiana

RESULT 4

tissues that include callus, roots from liquid culture
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
challenged leaves (Pseudomonas syringae pv tomato 12 hr;
Xanthomonas campestris pv campestris 12 hr, 18hr;
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
campestris pv vesicatoria 18hr). RNA was isolated from
these tissues and pooled in approximately equal molar
amounts."

ORIGIN

Query Match 15 78; Score 812.4; DB 7; Length 910;
Best Local Similarity 99.9%; Pred. No. 5.3e-211; Indels 0; Gaps 0;
Matches 813; Conservative 0; Mismatches 1;
Qy 3729 GATTGAACAAGATGGATTGCAACGAGCAATCGGCTGCTCTGATGCCCGCTTGGTGGAGAGGCTATTTCGG 3788
Db 4 GATTGAACAAGATGGATTGCAACGAGCAATCGGCTGCTCTGATGCCCGCTTGGTGGAGAGGCTATTTCGG 63
Qy 3789 CTATGACTGGGCAACAGACAAATCGGCTGCTCTGATGCCCGCTTGGTGGAGAGGCTGTCAGC 3848
Db 64 CTATGACTGGGCAACAGACAAATCGGCTGCTCTGATGCCCGCTTGGTGGAGAGGCTGTCAGC 123
Qy 3849 GCAGGGGGCGCGGCTCTTTTGTCAAGACGACCTGTCGGTGCCTGATGAATGAATGCA 3908
Db 124 GCAGGGGGCGCGGCTCTTTTGTCAAGACGACCTGTCGGTGCCTGATGAATGAATGCA 183
Qy 3909 GGACGAGCGAGCGCGGCTATGCTGGCTGGCCACACGCGGCGTTCCTTTCGCGAGCTGTGCT 3968
Db 184 GGACGAGCGAGCGCGGCTATGCTGGCTGGCCACACGCGGCGTTCCTTTCGCGAGCTGTGCT 243
Qy 3969 CGACGTTGTCACTGAAGCGGGAAGGAGCTGGCTCTATTGGGCGAAGTCCCGGGCAGGA 4028
Db 244 CGACGTTGTCACTGAAGCGGGAAGGAGCTGGCTCTATTGGGCGAAGTCCCGGGCAGGA 303
Qy 4029 TCTCTGCTCATCTCACCTTGTCTGCGGAGAAAGTATCCATCATGCTGATGATCAATGCG 4088
Db 304 TCTCTGCTCATCTCACCTTGTCTGCGGAGAAAGTATCCATCATGCTGATGATCAATGCG 363
Qy 4089 GCGGCTGCATACGCTTGCATCGGCTACTGCGCCATTTCGACCAACGAGCAACATCGCAT 4148
Db 364 GCGGCTGCATACGCTTGCATCGGCTACTGCGCCATTTCGACCAACGAGCAACATCGCAT 423
Qy 4149 CGAGCGAGCAGTACTCTCGGATGGAAGCGGCTTTGTCGATCAGGATGATCTGGACGAAGA 4208
Db 424 CGAGCGAGCAGTACTCTCGGATGGAAGCGGCTTTGTCGATCAGGATGATCTGGACGAAGA 483
Qy 4209 GCATCAGGGGCTCGCGCAGCGCAACTGTTCCGCGAGGCTCAAGCGCGCATGCCCGACGG 4268
Db 484 GCATCAGGGGCTCGCGCAGCGCAACTGTTCCGCGAGGCTCAAGCGCGCATGCCCGACGG 543
Qy 4269 CGAGGATCTCGTCTGACCCATGCGGATGCTGCTGCGGAATATCATGTGGAAAAATGG 4328
Db 544 CGAGGATCTCGTCTGACCCATGCGGATGCTGCTGCGGAATATCATGTGGAAAAATGG 603
Qy 4329 CCGCTTTTCTGGATTTCATCGACTGTGGCCGCGCTGGGTGTGGCGGACCGCTATCAGGACAT 4388
Db 604 CCGCTTTTCTGGATTTCATCGACTGTGGCCGCGCTGGGTGTGGCGGACCGCTATCAGGACAT 663
Qy 4389 AGCGTTGGCTACCGCTGATATTCGTAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCT 4448
Db 664 AGCGTTGGCTACCGCTGATATTCGTAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCT 723
Qy 4449 CGTGTCTTACGGTATCGCGCTCCGATTTCGAGCGCATCGCCTTCTATCGCCTTCTGA 4508
Db 724 CGTGTCTTACGGTATCGCGCTCCGATTTCGAGCGCATCGCCTTCTATCGCCTTCTGA 783
Qy 4509 CGAGTTCTTCGAGCGGAGCTCTGGGGTTCGATA 4542
Db 784 CGAGTTCTTCGAGCGGAGCTCTGGGGTTCGATA 817
|||||

```

CK291799
LOCUS      CK291799          933 bp      mRNA      linear      EST 02-AUG-2004
DEFINITION CK291799 Nicotiana benthamiana mixed tissue cDNA library,
normalized, full-length Nicotiana benthamiana cDNA clone NBMC477 5'
end, mRNA sequence.
ACCESSION  CK291799
VERSION     CK291799.1 GI:39872608
KEYWORDS    EST.
SOURCE      Nicotiana benthamiana
ORGANISM    Nicotiana benthamiana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1 (bases 1 to 933)
AUTHORS     Buell C.R., Hart A., Zismann V., Karamycheva, S.A., Day, B.,
Staskawicz, B., Jin, H. and Baker, B.
TITLE       Generation of EST sequences from Nicotiana benthamiana
JOURNAL     Unpublished (2003)
COMMENT     Other ESTs: EST754514
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/.
Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES             Location/Qualifiers
     source           1..933
                     /organism="Nicotiana benthamiana"
                     /mol_type="mRNA"
                     /db_xref="taxon:4100"
                     /clone="NBMC477"
                     /tissue_type="abiotic and biotic stress-treated leaves,
                     callus tissue and root tissue"
                     /lab_hosts="DH10B-Tona"
                     /clone_lib="Nicotiana benthamiana mixed tissue cDNA
                     library, normalized, full-length"
                     /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
                     supplier: RNA was isolated from Nicotiana benthamiana
                     tissues that include callus, roots from liquid culture
                     grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
                     cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
                     challenged leaves (Pseudomonas syringae pv tomato 12 hr;
                     Xanthomonas campestris pv campestris 12 hr, 18hr;
                     Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
                     campestris pv vesicatoria 18hr). RNA was isolated from
                     these tissues and pooled in approximately equal molar
                     amounts."
ORIGIN
Query Match      15.7%; Score 812.4; DB 7; Length 933;
Best Local Similarity 99.9%; Pred. No. 5.3e-211;
Matches 813; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      3729 GATTGAACAAGATGGATTGCACGAGGTTCTCCGGCCGCTTGGGTGCAGAGGCTATTTCGG 3788
Db      |||||||
Qy      3789 CTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCCGCGCTGTTCGGCTGTCCAGC 3848
Db      |||||||
Qy      106 CTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCCGCGCTGTTCGGCTGTCCAGC 165
Db      |||||||
Qy      3849 GCAGGGCGCCCGCTCTTTTGTTCAGACCGACTGTCGGTGCCCTCGATGAACTGCA 3908
Db      |||||||
Qy      166 GCAGGGCGCCCGCTCTTTTGTTCAGACCGACTGTCGGTGCCCTCGATGAACTGCA 225
Db      |||||||
Qy      3909 GGACGAGCGACGCGGCTATCGTGGCTGGCCACGACGGCGGCTTCTTTCGGCAGCTGTGCT 3968
Db      |||||||
Qy      226 GGAAGGCGACGCGGCTATCGTGGCTGGCCACGACGGCGGCTTCTTTCGGCAGCTGTGCT 285
Db      |||||||
Qy      3969 CGACGTTGTCTACTGAAGCGGAAGGACTGCTGCTCTATTGGCGCAAGTCCGGGCGCAGGA 4028
Db      |||||||
Qy      286 CGACGTTGTCTACTGAAGCGGAAGGACTGCTGCTCTATTGGCGCAAGTCCGGGCGCAGGA 345
Db      |||||||

```

```

Qy      4029 TCTCCTGTCTCATCTCACCTTGTCTCTGCTCCGAGAAAGTATCCATCATGCTGATGCAATGCG 4088
Db      |||||||
Qy      346 TCTCCTGTCTCATCTCACCTTGTCTCTGCTCCGAGAAAGTATCCATCATGCTGATGCAATGCG 405
Db      |||||||
Qy      4089 GCGGCTGCATACGCTTGATCCGGCTACTCTGCCCATTCGACCAACCAAGCGAAACATCGCAT 4148
Db      |||||||
Qy      406 GCGGCTGCATACGCTTGATCCGGCTACTCTGCCCATTCGACCAACCAAGCGAAACATCGCAT 465
Db      |||||||
Qy      4149 CGAGCGAGCAGCTACTCGGATGGAGCCGGTCTTGTGATCAGGATCATCTGGACGAGA 4208
Db      |||||||
Qy      466 CGAGCGAGCAGCTACTCGGATGGAGCCGGTCTTGTGATCAGGATCATCTGGACGAGA 525
Db      |||||||
Qy      4209 GCATCAGGGGCTCGCGCAGCGAACTGTTGCGCAGGCTCAAGCGCGCATGCCCGACGG 4268
Db      |||||||
Qy      526 GCATCAGGGGCTCGCGCAGCGAACTGTTGCGCAGGCTCAAGCGCGCATGCCCGACGG 585
Db      |||||||
Qy      4269 CGAGGATCTCGTCTGATGCCCATCGCGATGCTGCTGCCGAATATCATGTGGAAATGG 4328
Db      |||||||
Qy      586 CGAGGATCTCGTCTGATGCCCATCGCGATGCTGCTGCCGAATATCATGTGGAAATGG 645
Db      |||||||
Qy      4329 CCGCTTTTCTGGATTCTCATCTGCTGCGCGGCTGGGTGTGGCGGACCGCTATCAGGACAT 4388
Db      |||||||
Qy      646 CCGCTTTTCTGGATTCTCATCTGCTGCGCGGCTGGGTGTGGCGGACCGCTATCAGGACAT 705
Db      |||||||
Qy      4389 AGCGTTGGCTACCGCTGATATTCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCT 4448
Db      |||||||
Qy      706 AGCGTTGGCTACCGCTGATATTCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCT 765
Db      |||||||
Qy      4449 CGTGTCTTACGGTATCGCGCTCCCGATTCGACGCGCATCGCCTTCTATCGCCTTCTTGA 4508
Db      |||||||
Qy      766 CGTGTCTTACGGTATCGCGCTCCCGATTCGACGCGCATCGCCTTCTATCGCCTTCTTGA 825
Db      |||||||
Qy      4509 CGAGTTCTTCTGACGCGGACTCTGGGGTTTCGATA 4542
Db      |||||||
Qy      826 CGAGTTCTTCTGACGCGGACTCTGGGGTTTCGATA 859
Db      |||||||

RESULT 5
CK256977
LOCUS      CK256977          936 bp      mRNA      linear      EST 30-JUL-2004
DEFINITION CK256977 potato callus cDNA library, normalized and full-length
Solanum tuberosum cDNA clone POC170 5' end, mRNA sequence.
ACCESSION  CK256977
VERSION     CK256977.1 GI:39813957
KEYWORDS    EST.
SOURCE      Solanum tuberosum (potato)
ORGANISM    Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE   1 (bases 1 to 936)
AUTHORS     Buell C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
TITLE       Generation of ESTs from potato callus tissue
JOURNAL     Unpublished (2003)
COMMENT     Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/.
Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES             Location/Qualifiers
     source           1..936
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Kennebec"
                     /db_xref="taxon:4113"
                     /clone="POC170"
                     /tissue_type="callus"
                     /lab_hosts="DH10B-Tona"
                     /clone_lib="potato callus cDNA library, normalized and
                     full-length"

```



```
Db 377 CGCGTGCATACGCTTGATCCGGCTACCTGCCATTCGACCAACCAAGCAATCGCAT 436
Qy 4149 CGAGGAGCAGTACTCGATGGAAGCCGGCTTCTGTCGATCAGATGATCTGGACGAAGA 4208
Db 437 CGAGCGAGCAGTACTCGATGGAAGCCGGCTTCTGTCGATCAGATGATCTGGACGAAGA 496
Qy 4209 GCATCAGGGGCTCGCGCAGCCGCAACTCTTCGCGAGGCTCAAGCGCGCATGCCGACGG 4268
Db 497 GCATCAGGGGCTCGCGCAGCCGCAACTCTTCGCGAGGCTCAAGCGCGCATGCCGACGG 556
Qy 4269 CGAGGATCTCGTGCAGCCATCGCGATGCGCTTGCCTGATCATGCTGGGAAATGG 4328
Db 557 CGAGGATCTCGTGCAGCCATCGCGATGCGCTTGCCTGATCATGCTGGGAAATGG 616
Qy 4329 CCGCTTTCTCGATTCATGACTGCGCGGCTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 4388
Db 617 CCGCTTTCTCGATTCATGACTGCGCGGCTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 676
Qy 4389 AGCGTTGGCTACCGCTGATATTGCTGAAGAGCTTGGGGGCAATGGGCTGACCGCTTCCT 4448
Db 677 AGCGTTGGCTACCGCTGATATTGCTGAAGAGCTTGGGGGCAATGGGCTGACCGCTTCCT 736
Qy 4449 CGTGCCTTACCGATTCGCGGCTCCCGATTCGAGCGCATCGCCTTCTATCGCCTTCTGA 4508
Db 737 CGTGCCTTACCGATTCGCGGCTCCCGATTCGAGCGCATCGCCTTCTATCGCCTTCTGA 796
Qy 4509 CGAGTTCTTCGAGCGGAGCTCGGGTTCGATA 4542
Db 797 CGAGTTCTTCGAGCGGAGCTCGGGTTCGAAA 830
```

RESULT 7

CK283361
LOCUS EST746083 Nicotiana benthamiana mixed tissue cDNA library,
DEFINITION normalized, full-length Nicotiana benthamiana cDNA clone NBWAG50 5'
end, mRNA sequence.

CK283361
VERSION CK283361.1 GI:3985898

KEYWORDS

SOURCE

ORGANISM

Nicotiana benthamiana
Nicotiana benthamiana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.

REFERENCE

AUTHORS

1 (bases 1 to 954)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,
Staskiewicz, B., Jin, H. and Baker, B.

TITLE

Generation of EST sequences from Nicotiana benthamiana

JOURNAL

Unpublished (2003)

COMMENT

Other ESTs: EST746084
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr. Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via <http://genome.arizona.edu/orders/>.
Seq primer: ANT TAG GTG ACA CTA TAG.

FEATURES

source

1. 954
Location/Qualifiers

/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clone="NBWAG50"
/tissue type="abiotic and biotic stress-treated leaves,
callus tissue and root tissue"
/lab host="DH10B-TonA"
/clone_lib="Nicotiana benthamiana mixed tissue cDNA
library, normalized, full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Nicotiana benthamiana
tissues that include callus, roots from liquid culture

grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
challenged leaves (Pseudomonas syringae pv tomato 12 hr;
Xanthomonas campestris pv phaseolicola 18hr, 18hr;
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
campestris pv vesicatoria 18hr). RNA was isolated from
these tissues and pooled in approximately equal molar
amounts."

ORIGIN

Query Match 15.7%; Score 812.4; DB 7; Length 954;
Best Local Similarity 99.9%; Pred. No. 5.3e-211;
Matches 813; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 3729 GATTGAACAGATGGATTGACGACAGGTTCTCCGGCCGCTTGGGTGAGAGGCTATTTCGG 3788
Db 14 GATTGAACAGATGGATTGACGACAGGTTCTCCGGCCGCTTGGGTGAGAGGCTATTTCGG 73
Qy 3789 CTATGACTGGGCACAAACAGACAAATCGGCTGCTCTGATGCGCGCCGCTGTTCCGGCTGTGAGC 3848
Db 74 CTATGACTGGGCACAAACAGACAAATCGGCTGCTCTGATGCGCGCCGCTGTTCCGGCTGTGAGC 133
Qy 3849 GCAGGGGCGCCCGGCTCTTTTGTCAAGACCGACCTCTCGGTGCTCCGCTGCTGTAATGAATGCA 3908
Db 134 GCAGGGGCGCCCGGCTCTTTTGTCAAGACCGACCTCTCGGTGCTCCGCTGCTGTAATGAATGCA 193
Qy 3909 GGAGGAGGACGCGGCTATCGTGGCTGGCCACGACGCGGCGTTCCTTGGCAGCTGTGCT 3968
Db 194 GGAGGAGGACGCGGCTATCGTGGCTGGCCACGACGCGGCGTTCCTTGGCAGCTGTGCT 253
Qy 3969 CGACGTTGTCACTGAAGCGGGAAGGAGTGGCTGCTATTGGGGCAAGTGGCGGGGACGGA 4028
Db 254 CGACGTTGTCACTGAAGCGGGAAGGAGTGGCTGCTATTGGGGCAAGTGGCGGGGACGGA 313
Qy 4029 TCTCTGTCTATCTCACCTTCTCTCGCGAAGATATCCATCATATGCTGATGCAATGCG 4088
Db 314 TCTCTGTCTATCTCACCTTCTCTCGCGAAGATATCCATCATATGCTGATGCAATGCG 373
Qy 4089 CGCGCTGCATACGCTTGATCGGCTACCTGCCATTCGACCCACCAAGCGGAACATCGCAT 4148
Db 374 CGCGCTGCATACGCTTGATCGGCTACCTGCCATTCGACCCACCAAGCGGAACATCGCAT 433
Qy 4149 CGAGCGAGCAGCTACTCGGATGGAAGCGGCTTCTGTCGATCAGGATGATCTGGACGAAGA 4208
Db 434 CGAGCGAGCAGCTACTCGGATGGAAGCGGCTTCTGTCGATCAGGATGATCTGGACGAAGA 493
Qy 4209 GCATCAGGGGCTCGCGCAGCCGAACTGTTTCGCCAGGCTCAAGCGCGCATGCCGACGG 4268
Db 494 GCATCAGGGGCTCGCGCAGCCGAACTGTTTCGCCAGGCTCAAGCGCGCATGCCGACGG 553
Qy 4269 CGAGGATCTCGTGCATGCCATGCGGATGCTGCTGCCGATATCATGTTGGGAAATGG 4328
Db 554 CGAGGATCTCGTGCATGCCATGCGGATGCTGCTGCCGATATCATGTTGGGAAATGG 613
Qy 4329 CCGCTTTCTGGATTCTATCGACTGTGGCGGGCTGGGTGGGTGGGTGGGTGGGTGGGTGG 4388
Db 614 CCGCTTTCTGGATTCTATCGACTGTGGCGGGCTGGGTGGGTGGGTGGGTGGGTGGGTGG 673
Qy 4389 AGCGTTGGCTACCGCTGATATTGCTGAAGAGCTTGGCGGCGAAATGGGCTGACCGCTTCCT 4448
Db 674 AGCGTTGGCTACCGCTGATATTGCTGAAGAGCTTGGCGGCGAAATGGGCTGACCGCTTCCT 733
Qy 4449 CGTGCCTTACGGTATCGCGCTCCCGATTCGAGCGCATCGCCTTCTATCGCCTTCTGA 4508
Db 734 CGTGCCTTACGGTATCGCGCTCCCGATTCGAGCGCATCGCCTTCTATCGCCTTCTGA 793
Qy 4509 CGAGTTCTTCGAGCGGAGCTCTGGGGTTCGATA 4542
Db 794 CGAGTTCTTCGAGCGGAGCTCTGGGGTTCGAAA 827

RESULT 8
CK291519

LOCUS CK291519 804 bp mRNA linear EST 02-AUG-2004
 DEFINITION EST754233 Nicotiana benthamiana mixed tissue cDNA library,
 normalized, full-length Nicotiana benthamiana cDNA clone NEMC276 5',
 end, mRNA sequence.
 ACCESSION CK291519
 VERSION CK291519.1 GI:39872047
 KEYWORDS EST.
 SOURCE Nicotiana benthamiana
 ORGANISM Nicotiana benthamiana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Nicotiana.
 REFERENCE 1 (bases 1 to 804)
 AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A., Day,B.,
 Staskawicz,B., Jin,H. and Baker,B.
 TITLE Generation of EST sequences from Nicotiana benthamiana
 JOURNAL Unpublished (2003)
 COMMENT Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from the University of Arizona Genomics
 Institute via <http://genome.arizona.edu/orders/>.
 Seq primer: ATT TAG GTG ACA CTA TAG.
 FEATURES
 Location/Qualifiers
 1..804
 /organism="Nicotiana benthamiana"
 /mol_type="mRNA"
 /db_xref="taxon:4100"
 /clone="NEMC276"
 /tissue_type="abiotic and biotic stress-treated leaves,
 callus tissue and root tissue"
 /lab_host="DH10B-Tona"
 /clone_lib="Nicotiana benthamiana mixed tissue cDNA
 library, normalized, full-length"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
 supplier: RNA was isolated from Nicotiana benthamiana
 tissues that include callus, roots from liquid culture
 grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
 cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
 challenged leaves (Pseudomonas syringae pv tomato 12 hr;
 Xanthomonas campestris pv campestris 12 hr, 18hr;
 Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
 campestris pv vesicatoria 18hr). RNA was isolated from
 these tissues and pooled in approximately equal molar
 amounts."
 ORIGIN
 Query Match 15.5%; Score 803; DB 7; Length 804;
 Best Local Similarity 100.0%; Pred. No. 1.9e-208;
 Matches 803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3729 GATTGACAAAGATGATGACGAGGTTCTCCGCGCTGGGTGGAGAGGCTATTCCG 3788
 DB 2 GATTGACAAAGATGATGACGAGGTTCTCCGCGCTGGGTGGAGAGGCTATTCCG 61
 QY 3789 CTATGACTGGGCAACACAGCAATCGGCTGCTGATGCCCGCTGTTCGGCTGTGAGC 3848
 DB 62 CTATGACTGGGCAACACAGCAATCGGCTGCTGATGCCCGCTGTTCGGCTGTGAGC 121
 QY 3849 CGAGGGGCGCGGTTCTTTTGTCAAGACCGACTGTCGGTCCCTGAAATGCA 3908
 DB 122 CGAGGGGCGCGGTTCTTTTGTCAAGACCGACTGTCGGTCCCTGAAATGCA 181
 QY 3909 GGAGGAGCGCGGCTATCGTGGCTGGCCACGCGGCTTCTTGGCGAGCTGTGCT 3968
 DB 182 GGAGGAGCGCGGCTATCGTGGCTGGCCACGCGGCTTCTTGGCGAGCTGTGCT 241
 QY 3969 CGAGCTTCTCAAGCGGGAAGGACTGCTGCTATTGGCGGAAGTGGCGGCGGAGGA 4028
 DB 242 CGAGCTTCTCAAGCGGGAAGGACTGCTGCTATTGGCGGAAGTGGCGGCGGAGGA 301
 QY 4029 TCTCCTGTCTATCTACCTTGTCTCTGCGGAGAAAGTATCCATCATGCGTGAATGCG 4088

Db 302 TCTCCTGTCTATCTACCTTGTCTTCTGCGGAGAAAGTATCCATCATGGTGTGATGCAATCGG 361
 QY 4089 GCGGCTGCATACGCTTGTATCGGCTACTCTGCCATTCGACCACCAAGCAACATCGCAT 4148
 Db 362 GCGGCTGCATACGCTTGTATCGGCTACTCTGCCATTCGACCACCAAGCAACATCGCAT 421
 QY 4149 CGAGCGAGCGTACTCGATGGAAGCGGCTTGTTCGATCAGATGATCTGGAGCAAGA 4208
 Db 422 CGAGCGAGCGTACTCGATGGAAGCGGCTTGTTCGATCAGATGATCTGGAGCAAGA 481
 QY 4209 GCATCAGGGGCTCGCGCAGCGCAACTTTCGCCAGGCTCAAGCGGCGCATGCCGACGG 4268
 Db 482 GCATCAGGGGCTCGCGCAGCGCAACTTTCGCCAGGCTCAAGCGGCGCATGCCGACGG 541
 QY 4269 CGAGGATCTCGTCTGACCCATGCGGCTGCTTCCGAATATCATGTGGAAAAATGG 4328
 Db 542 CGAGGATCTCGTCTGACCCATGCGGCTGCTTCCGAATATCATGTGGAAAAATGG 601
 QY 4329 CCGCTTTTCTGGATTCATCGACTGTGCGCGGCTGGGTGGCGGACCGCTATCAGGACAT 4388
 Db 602 CCGCTTTTCTGGATTCATCGACTGTGCGCGGCTGGGTGGCGGACCGCTATCAGGACAT 661
 QY 4389 AGCGTTGCTACCGCTGATATTCGTAAGAGCTTGGCGGCGCAATGGGCTGACCGCTTCT 4448
 Db 662 AGCGTTGCTACCGCTGATATTCGTAAGAGCTTGGCGGCGCAATGGGCTGACCGCTTCT 721
 QY 4449 CGTCTTTTACGGTATCGCGCTCCCGATTCGACGCGATTCGCTTCTATCGCTTCTTGA 4508
 Db 722 CGTCTTTTACGGTATCGCGCTCCCGATTCGACGCGATTCGCTTCTATCGCTTCTTGA 781
 QY 4509 CGAGTTCTTCTGAGCGGAGCTCT 4531
 Db 782 CGAGTTCTTCTGAGCGGAGCTCT 804
 RESULT 9
 CK288185
 LOCUS
 DEFINITION EST750907 Nicotiana benthamiana mixed tissue cDNA library,
 normalized, full-length Nicotiana benthamiana cDNA clone NEMBE61 5',
 end, mRNA sequence.
 ACCSSION CK288185
 VERSION CK288185.1 GI:39865462
 KEYWORDS EST.
 SOURCE Nicotiana benthamiana
 ORGANISM Nicotiana benthamiana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Nicotiana.
 REFERENCE 1 (bases 1 to 811)
 AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A., Day,B.,
 Staskawicz,B., Jin,H. and Baker,B.
 TITLE Generation of EST sequences from Nicotiana benthamiana
 JOURNAL Unpublished (2003)
 COMMENT Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from the University of Arizona Genomics
 Institute via <http://genome.arizona.edu/orders/>.
 Seq primer: ATT TAG GTG ACA CTA TAG.
 FEATURES
 Location/Qualifiers
 1..811
 /organism="Nicotiana benthamiana"
 /mol_type="mRNA"
 /db_xref="taxon:4100"
 /clone="NEMBE61"
 /tissue_type="abiotic and biotic stress-treated leaves,
 callus tissue and root tissue"
 /lab_host="DH10B-Tona"
 /clone_lib="Nicotiana benthamiana mixed tissue cDNA
 library, normalized, full-length"

/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Nicotiana benthamiana tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (Pseudomonas syringae pv tomatato 12 hr; Xanthomonas campestris pv campestris 12 hr, 18hr; Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas campestris pv vesicatoria 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."

```
ORIGIN
Query Match      15.3%; Score 794; DB 7; Length 811;
Best Local Similarity 99.4%; Pred. No. 5.8e-206;
Matches 797; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3706 ACAGATGAGGATCGTTTCGATGATTGAACAAGATGATTCACGACGAGGTCTCCGGCC 3765
Db 10 ACCGGATCTGGATCGTTTCGATGATTGAACAAGATGATTCACGACGAGGTCTCCGGCC 69

Qy 3766 GCTTGGTGGAGAGGCTATTGCGGTATGACTGGGCAACAAGACAACTCGGCTCTCGAT 3825
Db 70 GCTTGGTGGAGAGGCTATTGCGGTATGACTGGGCAACAAGACAACTCGGCTCTCGAT 129

Qy 3826 GCGCGCGTTCGCGCTCTGACGCGGCGCGCGGTTCTTTTGTCAAGACGACCTG 3885
Db 130 GCGCGCGTTCGCGCTCTGACGCGGCGCGCGGTTCTTTTGTCAAGACGACCTG 189

Qy 3886 TCCGCTGCGCTGAATGAATCTGAGACGAGGCGCGCTATCGTCTGGCCACGACG 3945
Db 190 TCCGCTGCGCTGAATGAATCTGAGACGAGGCGCGCTATCGTCTGGCCACGACG 249

Qy 3946 GCGCTTCTTCGCGAGTGTCTGCACTTGTCTGAAAGCGGAAAGGACTGGCTGCTA 4005
Db 250 GCGCTTCTTCGCGAGTGTCTGCACTTGTCTGAAAGCGGAAAGGACTGGCTGCTA 309

Qy 4006 TTGGCGAAGTCCGCGGCGAGGATCTCTGTCATCTCACTTGTCTTCTGCGGAAAGTA 4065
Db 310 TTGGCGAAGTCCGCGGCGAGGATCTCTGTCATCTCACTTGTCTTCTGCGGAAAGTA 369

Qy 4066 TCCATCATGCTGATGCAATCGCGGCTGCATACGCTTGATCGGGTACCTGCGCCATTC 4125
Db 370 TCCATCATGCTGATGCAATCGCGGCTGCATACGCTTGATCGGGTACCTGCGCCATTC 429

Qy 4126 GACCAACGAGAAACATCGCATCGAGCGACGCTACTCGGATGGAAGCGGCTTGTGTC 4185
Db 430 GACCAACGAGAAACATCGCATCGAGCGACGCTACTCGGATGGAAGCGGCTTGTGTC 489

Qy 4186 GATCAGGATGATCTGGACGAGACATCAGGGGCTCGCGGCGGACGCTGTTCCGCCAGG 4245
Db 490 GATCAGGATGATCTGGACGAGACATCAGGGGCTCGCGGCGGACGCTGTTCCGCCAGG 549

Qy 4246 CTCAAGCGCGCATGCCCGGCGAGGATCTCGTCTGACCCATGCGGATGCGCTGCTTG 4305
Db 550 CTCAAGCGCGCATGCCCGGCGAGGATCTCGTCTGACCCATGCGGATGCGCTGCTTG 609

Qy 4306 CCGAATATCATGGTGGAAATAGCGCGCTTTTCTGATTCATCGACTGTGGCGCGCTGGGT 4365
Db 610 CCGAATATCATGGTGGAAATAGCGCGCTTTTCTGATTCATCGACTGTGGCGCGCTGGGT 669

Qy 4366 GTGGCGGACCCCTATCAGGACATAGCGTTGGCTACCGGTGATATGCTGAAGACTTGGC 4425
Db 670 GTGGCGGACCCCTATCAGGACATAGCGTTGGCTACCGGTGATATGCTGAAGACTTGGC 729

Qy 4426 GCGAATGGGCTGACCGCTTCTCTGCTTTTACGCTATCGCGCTTCCCGATTTCGACGCG 4485
Db 730 GCGAATGGGCTGACCGCTTCTCTGCTTTTACGCTATCGCGCTTCCCGATTTCGACGCG 789

Qy 4486 ATCGCCTTCTATCGCCTTTCTTG 4507
Db 790 ATCGCCTTCTATCGCCTTTCTTG 811
```

RESULT 10

DN548862 LOCUS 1407127 MARC 7BOV Bos taurus cDNA 5', mRNA linear EST 11-MAR-2005

DEFINITION DN548862

ACCESSION DN548862

VERSION DN548862.1 GI:61013550

KEYWORDS EST.

SOURCE Bos taurus (cow)

ORGANISM Bos taurus

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@mail.marc.usda.gov

Single pass sequencing. Bases called with phred v0.020425.c and

trimmed with the aid of the trim_alt option. Vector identified with

cross_match v0.990329.

Plate: OQP8005 row: C column: 19

Seq primer: GTAATACGACTCATATAGG.

FEATURES

source

1..793

/organism="Bos taurus"

/mol_type="mRNA"

/db_xref="taxon:9913"

/tissue_type="pooled"

/lab_host="DH10B"

/clone_lib="MARC 7BOV"

/notes="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;

Library made with RNA pooled from multiple tissues

including ovary, hindbrain, uterus, and day-30 whole

embryos."

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 1.1e-205;

Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1036 TAAATCGGTCTGCGGTCTTTAGGACATATGAAAGTATGGCACAGTGGGATGACCTTCTG 1095

Db 1 TAAATCGGTCTGCGGTCTTTAGGACATATGAAAGTATGGCACAGTGGGATGACCTTCTG 60

Qy 1096 ATCAGCAAGAGGACATGACAGCTGTACAGAGTCTGTGAAAGTTCGATGCTCGCTCAGTGA 1155

Db 61 ATCAGCAAGAGGACATGACAGCTGTACAGAGTCTGTGAAAGTTCGATGCTCGCTCAGTGA 120

Qy 1156 CAGCTTGTCTCTCCCATCTCCCACTCTTAAATGCGCCCACTCTTCAAGAGAGTGAAGTCTT 1215

Db 121 CAGCTTGTCTCTCCCATCTCTTAAATGCGCCCACTCTTCAAGAGAGTGAAGTCTT 180

Qy 1216 ATAAACTGCACTGATCACCCTTTATCTCATTTGTTGTAGTTCTCGTCCCATCATTTG 1275

Db 181 ATAAACTGCACTGATCACCCTTTATCTCATTTGTTGTAGTTCTCGTCCCATCATTTG 240

Qy 1276 GCATAGTGGCAGTCACTGCTCTGAAATGGGAAACGAAGAATTCACGGTTGGCTCAGTTA 1335

Db 241 GCATAGTGGCAGTCACTGCTCTGAAATGGGAAACGAAGAATTCACGGTTGGCTCAGTTA 300

Qy 1336 ATCAGATATATCTCCAGTCCCGAGGCAAGGAATGCGAGTGAAGATGAATAGAT 1395

Db 301 ATCAGATATATCTCCAGTCCCGAGGCAAGGAATGCGAGTGAAGATGAATAGAT 360

Qy 1396 TTCGAGAGGCTGTGATGGAACGATGAGCAACATGGAAGCAGAAATCCAGTATCTTTTCAG 1455

Db 361 TTCGAGAGCTGTGATGACCGCATGAGCAACATGGAAGCAGATCCAGTATCTTTTCAG 420

Qy 1456 ATAATGAAGCCAAATCTCTAGATGCTAAGATTTCCAAAATTTCCAGCAATAACAACTGATC 1515

Db 421 ATAATGAAGCCAAATCTCTAGATGCTAAGATTTCCAAAATTTCCAGCAATAACAACTGATC 480

Qy 1516 AAAGATTAAATGATGTTCTTTTCCAGCTAAATTCCTTACTTTCTCCATCCAGCAACATG 1575

Db 481 AAAGATTAAATGATGTTCTTTTCCAGCTAAATTCCTTACTTTCTCCATCCAGCAACATG 540

Qy 1576 AGAATATCATAGGGGATATCTCCAAGTCATTAGTAGGTCTGAACACCAACAGTACTTGATT 1635

Db 541 AGAATATCATAGGGGATATCTCCAAGTCATTAGTAGGTCTGAACACCAACAGTACTTGATT 600

Qy 1636 TGCAGTTCAGTATTGAACACATGTAATGCGAGTCCCAAGAGAAATGCAATTTAAACAACAG 1695

Db 601 TGCAGTTCAGTATTGAACACATGTAATGCGAGTCCCAAGAGAAATGCAATTTAAACAACAG 660

Qy 1696 AGGAGATCGTAAATTAGAGAGCGTATATACAATGATCAGCAGAAATTAAGTCTCTAG 1755

Db 661 AGGAGATCGTAAATTAGAGAGCGTATATACAATGATCAGCAGAAATTAAGTCTCTAG 720

Qy 1756 ATGAAAACAAGTATATTGTGAACAGGAAATAAAGGGGAAATGAACTGTTGAATAATA 1815

Db 721 ATGAAAACAAGTATATTGTGAACAGGAAATAAAGGGGAAATGAACTGTTGAATAATA 780

Qy 1816 TCACTAATGATCT 1828

Db 781 TCACTAATGATCT 793

RESULT 11

DR696294

LOCUS

DEFINITION SMV3-0053 Atlantic Salmon macrophage - Aeromonas in vivo infection

ACCESSION

DR696294

VERSION

EST.

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei;

Protacanthopterygii; Salmoniformes; Salmonidae; Salmo;

1 (bases 1 to 767)

AUTHORS

Douglas, S.E., Tsou, S.C.M., Penny, S., Melville, K., Liebscher, R. and

Ewart, K.V.

TITLE

Unpublished (2005), Douglas, S.E., et al

JOURNAL

Unpublished (2005)

COMMENT

Contact: Douglas S

Genome Sciences

NRC Institute for Marine Biosciences

1411 Oxford St., Halifax, Nova Scotia B3H3Z1 Canada

Tel: (902) 426-4991

Fax: (902) 426-9413

Email: susan.douglas@nrc.ca

Seq primer: M13R.

Location/Qualifiers

1..767

/organism="Salmo salar"

/mol_type="mRNA"

/db_xref="taxon:8030"

/clone="SMV3-0053"

/cell_type="enriched macrophages"

/dev_stage="adult"

/clone_lib="Atlantic Salmon macrophage - Aeromonas in vivo

infection"

/note="Vector: pT-Adv; Libraries (SMV2, SMV3, SMV6, SMV7)

made by forward and reverse suppression subtractive

hybridization using salmon macrophage infected in vitro

for 0.5h (SMV2 and SMV3) or 2h (SMV6 and SMV7) with

Aeromonas salmonicida cultured in the intraperitoneal

space of fish vs. uninfected macrophages"

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches 738; Conservative

14.2%; Score 737.4; DB 8; Length 767;

99.9%; Pred. No. 2e-190;

0; Mismatches 1; Indels 0; Gaps 0;

Qy 3783 ATTGGCTATGACTGGGCAACAGACAAATCGGCTGCTCTGATCGCGCGGTTCGGCT 3842

Db 12 ATTGGCTATGACTGGGCAACAGACAAATCGGCTGCTCTGATCGCGCGGTTCGGCT 71

Qy 3843 GTACAGCGAGGGCGCGCGGTTCCTTTTGTCAAGACCGACCTGTCGGTCCCTGGAATGA 3902

Db 72 GTACAGCGAGGGCGCGCGGTTCCTTTTGTCAAGACCGACCTGTCGGTCCCTGGAATGA 131

Qy 3903 ACTGACGACGAGGACGCGGCTATCGTGGCTGGCCACGACGCGGCGTTCCTTCGCGAGC 3962

Db 132 ACTGACGACGAGGACGCGGCTATCGTGGCTGGCCACGACGCGGCGTTCCTTCGCGAGC 191

Qy 3963 TGTGCTCGAGCTTGTCTACTGAGCGGGAAGGACTGGCTGCTATTGGGCGAAGTGC CGGG 4022

Db 192 TGTGCTCGAGCTTGTCTACTGAGCGGGAAGGACTGGCTGCTATTGGGCGAAGTGC CGGG 251

Qy 4023 GCAGGATCTCTGTCATCTCACCTTGTCTCCGCGAGAAAGTATCCATCATGGCTGATGC 4082

Db 252 GCAGGATCTCTGTCATCTCACCTTGTCTCCGCGAGAAAGTATCCATCATGGCTGATGC 311

Qy 4083 AATGCGGCGGCTGCATACGCTTGTATCGGCTTACCTCGCCATTTCGACCAACGAGGAAACA 4142

Db 312 AATGCGGCGGCTGCATACGCTTGTATCGGCTTACCTCGCCATTTCGACCAACGAGGAAACA 371

Qy 4143 TCGCATCGAGCGACAGCTACTCGGATGGAAGCGGCTTGTTCGATCAGGATGATCTGGA 4202

Db 372 TCGCATCGAGCGACAGCTACTCGGATGGAAGCGGCTTGTTCGATCAGGATGATCTGGA 431

Qy 4203 CGAAGACATCAGGGGCTCGCGCCAGCGAACTGTTCCGCGAGGCTCAAGGCGCGATGCC 4262

Db 432 CGAAGACATCAGGGGCTCGCGCCAGCGAACTGTTCCGCGAGGCTCAAGGCGCGATGCC 491

Qy 4263 CGAAGACATCAGGGGCTCGCGCCAGCGAACTGTTCCGCGAGGCTCAAGGCGCGATGCC 4322

Db 492 CGAAGACATCAGGGGCTCGCGCCAGCGAACTGTTCCGCGAGGCTCAAGGCGCGATGCC 551

Qy 4323 AATGGCGGCTTTCGATTCATCGACTGTCGCGCGCTGGGTGCGGACCGCTATCA 4382

Db 552 AATGGCGGCTTTCGATTCATCGACTGTCGCGCGCTGGGTGCGGACCGCTATCA 611

Qy 4383 GGCATAGCTTGGCTACCGCTGATATGCTGAAGAGCTTGGCGGCGAATGGGCTGACCG 4442

Db 612 GGCATAGCTTGGCTACCGCTGATATGCTGAAGAGCTTGGCGGCGAATGGGCTGACCG 671

Qy 4443 CTTCTCTGCTTTTACCGTATCGCGCTTCCCGATTTCGAGCGCATCGCTTCTATCGCT 4502

Db 672 CTTCTCTGCTTTTACCGTATCGCGCTTCCCGATTTCGAGCGCATCGCTTCTATCGCT 731

Qy 4503 TCTTGACGAGTCTTTCTGA 4521

Db 732 TCTTGACGAGTCTTTCTGA 750

RESULT 12

CF409555

LOCUS

DEFINITION

CH3#061_D06MF Canine heart normalized cDNA Library in pluescript

CF409555

VERSION

CF409555.1

KEYWORDS

EST.

SOURCE

Canis familiaris (dog)

ORGANISM

Canis familiaris

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;

Canis.

REFERENCE

1 (bases 1 to 949)

949 bp mRNA linear EST 02-SEP-2003

Canis familiaris cDNA clone CH3#061_D06 5', mRNA sequence.

Canis familiaris

Canis familiaris

Canis familiaris

Canis familiaris

Canis familiaris

Canis familiaris

Canis familiaris

Canis familiaris

Canis familiaris

Canis familiaris

Canis familiaris

Canis familiaris

Canis familiaris

Canis familiaris

Canis familiaris

Canis familiaris

	ORIGIN	Query Match	Score	DB 10;	Length
Qy	1456 ATAATGAAGCCCAATCTCTCTAGATGCTAAGAAATTTCCAAAATTTTCAGCATCAACTGATC 1515	12.34;	97.34;	634.2;	837;
Db	590 ATAATGAAGCCCAATCTCTCTAGATGCTAAGAAATTTCCAAAATTTTCAGCATCAACTGATC 649	Best Local Similarity	Pred. No. 4.7e-162;		
Qy	1516 AAGAAATTTAATGATGTTCTTTTCCAGCTAAATTCCTTACTTTCTCTCCATCCAGGAACATG 1575	Matches 686; Conservative	0;	Mismatches 15;	Indels 4; Gaps 4;

RESULT 15

DR696687
 LOCUS SMV7-0178 Atlantic Salmon macrophage - Aeromonas in vivo infection EST 13-JUL-2005
 DEFINITION Salmo salar cDNA clone SMV7-0178, mRNA sequence.
 ACCESSION DR696687
 VERSION DR696687.1 GI:70787047
 KEYWORDS EST.
 SOURCE Salmo salar (Atlantic salmon)
 ORGANISM Salmo salar
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
 REFERENCE 1 (Bases 1 to 591)
 AUTHORS Douglas, S.E., Tsai, S.C.M., Penny, S., Melville, K., Liebscher, R. and Ewart, K.V.
 TITLE Unpublished (2005), Douglas, S.E., et al
 JOURNAL Unpublished (2005)
 COMMENT Contact: Douglas S
 Genome Sciences
 NRC Institute for Marine Biosciences
 1411 Oxford St., Halifax, Nova Scotia B3H3Z1 Canada
 Tel: (902) 426-4991
 Fax: (902) 426-9413
 Email: susan.douglas@nrc.ca
 Seq primer: M13R.

FEATURES
 source

1..591
 Location/Qualifiers
 /organism="Salmo salar"
 /mol_type="mRNA"
 /db_xref="taxon:8030"
 /clone="SMV7-0178"
 /cell_type="enriched macrophages"
 /dev_stage="adult"
 /clone_lib="Atlantic Salmon macrophage - Aeromonas in vivo infection"
 /note="Vector: pT-Adv; Libraries (SMV2, SMV3, SMV6, SMV7) made by forward and reverse suppression subtractive hybridization using salmon macrophage infected in vitro for 0.5h (SMV2 and SMV3) or 2h (SMV6 and SMV7) with Aeromonas salmonicida cultured in the intraperitoneal space of fish vs. uninfected macrophages"

ORIGIN

Query Match 11.4%; Score 589.4; DB 8; Length 591;
 Best Local Similarity 99.8%; Pred. No. 8.7e-150;
 Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 3812 TCGGCTGCTCTGATGCCCGCGCTGTTCCGGCTGTTCAGCGCAGGGGCGCGGCTCTTTTGG 3871
 Db 1 TCGGCTGCTCTGATGCCCGCGCTGTTCCGGCTGTTCAGCGCAGGGGCGCGGCTCTTTTGG 60
 Qy 3872 TCAAGACCGACTGTGCGGGTCCCTGAATGAATGCACTGCAGGACGAGCGCGGCTATCGT 3931
 Db 61 TCAAGACCGACTGTTCGGGTGCCCTGAATGAATGCACTGCAGGACGAGCGCGGCTATCGT 120
 Qy 3932 GCGTGGCCACGACGGCGGCTTCCTTTGCGCAGCTGTGCTGCAGCTTGTCACTGAAGCGGGAA 3991
 Db 121 GCGTGGCCACGACGGCGGCTTCCTTTGCGCAGCTGTGCTGCAGCTTGTCACTGAAGCGGGAA 180
 Qy 3992 GGGACTGGCTGCTATTGGGGCGAAGTGC CGGGGCGAGGATCTCTGTGATCTCACCTTGGTC 4051
 Db 181 GGGACTGGCTGCTATTGGGGCGAAGTGC CGGGGCGAGGATCTCTGTGATCTCCACCTTGGTC 240
 Qy 4052 CTGCGGAGAAAGATCCATCATGCTGATGCAATGCGCGGCTGCATACGCTTCATCCGG 4111
 Db 241 CTGCGGAGAAAGATCCATCATGCTGATGCAATGCGCGGCTGCATACGCTTCATCCGG 300
 Qy 4112 CTACCTGCCATTTCGACCACCAAGCGAAACATCGCATCGAGCGACGACTACTCGGATGG 4171
 Db 301 CTACCTGCCATTTCGACCACCAAGCGAAACATCGCATCGAGCGACGACTACTCGGATGG 360

Qy 4172 AAGCCGGTCTTGTCTGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCAGCCG 4231
 Db 361 AAGCCGGTCTTGTCTGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCAGCCG 420
 Qy 4232 AACTGTTTCGCCAGGCTCAAGCGCGGCATGCCGACGCGGAGGATCTCTGTCGTGACCCATG 4291
 Db 421 AACTGTTTCGCCAGGCTCAAGCGCGGCATGCCGACGCGGAGGATCTCTGTCGTGACCCATG 480
 Qy 4292 GCGATGCTCTGTTGCCGGAATATCATGGTGGAAAAATGSCCGCTTTTCTGGATTCTACACT 4351
 Db 481 GCGATGCTCTGTTGCCGGAATATCATGGTGGAAAAATGSCCGCTTTTCTGGATTCTACACT 540
 Qy 4352 GTGCGCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCC 4402
 Db 541 GTGCGCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCC 591

Search completed: February 27, 2006, 13:53:53
 Job time : 18624 secs

GenCore version 5.1.7
 Copyright (c) 1993 - 2006 Bioceleration Ltd.
 OM nucleic - nucleic search, using sw model
 Run on: February 27, 2006, 08:16:10 ; Search time 2790 Seconds
 (without alignments)
 12366.707 Million cell updates/sec

Title: US-10-618-570-1
 Perfect score: 5177
 Sequence: 1 ttgaaagaccaccaccgta.....agcgggggttttatttgg 5177

Scoring table: IDENTITY_NUC
 Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_21.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001as.*
- 5: Geneseq2001bs.*
- 6: Geneseq2002as.*
- 7: Geneseq2002bs.*
- 8: Geneseq2003as.*
- 9: Geneseq2003bs.*
- 10: Geneseq2003cs.*
- 11: Geneseq2003ds.*
- 12: Geneseq2004as.*
- 13: Geneseq2004bs.*
- 14: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5165.8	99.8	5177	2	Aaz09997
2	2077.2	40.1	7699	2	Aav81410
3	1972.6	38.1	7436	13	Adq06075
4	1916.6	37.0	7546	2	Aax90485
5	1661.6	32.1	8316	2	Aaq74450
6	1644.6	31.8	7165	2	Aax90483
7	1629.4	31.5	5594	2	Aax90482
8	1588.6	30.7	5130	6	Adm28311
9	1588.6	30.7	5130	12	Adm68972
10	1588.6	30.7	5130	14	Adw28107
11	1588.6	30.7	5130	14	Adw77874
12	1486	28.7	5874	2	Aax90484
13	1486	28.7	6141	2	Aax90481
14	1486	28.7	6522	2	Aax90482
15	1486	28.7	6620	6	Adm28311
16	1486	28.7	9115	2	Aaq74449
17	1484.4	28.7	6365	2	Aaq41173
18	1481	28.6	6046	12	Adg83262
19	1463	28.3	6825	14	Aeb48081

20	1463	28.3	6851	14	AEB48080
21	1453.6	28.1	5865	2	AAV04002
22	1448.4	28.0	8591	2	AAT84562
23	1448.4	28.0	8591	2	AAT84561
24	1448.4	28.0	8591	2	AAT87083
25	1448.4	28.0	8591	2	AAT87084
26	1448.4	28.0	8591	2	AAV04866
27	1448.4	28.0	8591	2	AAV04865
28	1448.4	28.0	8591	2	AAV05850
29	1448.4	28.0	8591	2	AAV05849
30	1448	28.0	6444	6	AAI67595
31	1441	27.8	7170	6	AAI67595
32	1434	27.7	6700	4	AAH76187
33	1434	27.7	8518	4	AAH74524
34	1434	27.7	8518	4	AAH76188
35	1432.4	27.7	6700	4	AAH74523
36	1429.8	27.6	8078	10	ADBS2213
37	1425	27.5	5689	2	AAZ11445
38	1425	27.5	5689	2	AAZ11445
39	1425	27.5	5689	3	AAA96208
40	1425	27.5	9849	3	AAA96221
41	1352	26.1	7925	12	ADP74702
42	1292.8	25.0	9830	3	AAA96220
43	1206.8	23.3	6505	9	ADA12886
44	1186.6	22.9	7185	12	ADP74699
45	1186.6	22.9	7185	12	ADP74701

ALIGNMENTS

RESULT 1

AAZ09997
 ID AAZ09997 standard; cDNA; 5177 BP.

XX AC AAZ09997;

XX DT 03-DEC-1999 (first entry)

DE DE Bovine scavenger receptor class A (SCR)/avidin fusion protein cDNA.

KW KW Scavenger receptor class A; SCR; avidin; fusion protein; bovine; ECD;
 KW membrane-spanning domain; extracellular domain; biotin-binding activity;
 KW endocytosis; ss.

XX OS Synthetic.

XX OS Bos taurus.

FH Key Location/Qualifiers

FT CDS 1071..2269

FT FT /*tag= a

FT FT /product= "SCR/avidin fusion protein"

XX WO9942577-A2.

XX PD 26-AUG-1999.

XX PF 23-FEB-1999; 99WO-GB0000546.

XX PR 23-FEB-1998; 98GB-00003757.

XX PR 24-JUN-1998; 98GB-00013653.

XX PA (EURO-) EUROGENE LTD.

XX PI Yla-Herttua S, Kulomaa M, Lehtolainen P, Marjomaki V, Airenne K;

XX DR WPI; 1999-561345/47.

XX DR P-PSDB; AAY03098.

XX PT New fusion proteins having an extracellular domain with biotin-binding
 PT activity, used to target biotinylated molecules to specific sites in
 XX tissues.

PS Disclosure; Page 15-21; 23pp; English.

CC This invention describes a novel protein (A) which comprises a membrane-
CC spanning domain and an extracellular domain (ECD), where the ECD
CC comprises biotin-binding activity. Using the proteins or encoding nucleic
CC acid molecules it is possible to target biotinylated molecules to
CC specific sites in tissues. Molecules targeted in this way may be taken up
CC by the tissues or cells by endocytosis, allowing the molecules to exert
CC their effects within or on the cell. This sequence encodes a bovine
CC scavenger receptor class A/avidin fusion protein which is used in the
CC description of the invention

XX Sequence 5177 BP; 1356 A; 1222 C; 1390 G; 1209 T; 0 U; 0 Other;

Query Match 99.8%; Score 5165.8; DB 2; Length 5177;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 5170; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TTTGAAAGACCCACCCCTAGGTGGCAAGCTAGCTTAAGTAACGCCACTTTGCAAGGCAT 60

DB 1 TTTGAAAGACCCACCCCTAGGTGGCAAGCTAGCTTAAGTAACGCCACTTTGCAAGGCAT 60

QY 61 GGAATAATACATACTAGTAATAGAAAAGTTTCAGATCAAGTTCAGGAACAAGAAAACAGC 120

DB 61 GGAATAATACATACTAGTAATAGAAAAGTTTCAGATCAAGTTCAGGAACAAGAAAACAGC 120

QY 121 TGAATACCAACAGGATATCTGTGGTAAGCGGTTCTGCGCCCGGCTCAGGCGCAAGAAC 180

DB 121 TGAATACCAACAGGATATCTGTGGTAAGCGGTTCTGCGCCCGGCTCAGGCGCAAGAAC 180

QY 181 GATGAGACAGCTGAGTGGCGCAACAGGATATCTGTGGTAAGCGGTTCTGCGCCCGG 240

DB 181 GATGAGACAGCTGAGTGGCGCAACAGGATATCTGTGGTAAGCGGTTCTGCGCCCGG 240

QY 241 CTCGGGGCAAGAACAGATGTCCTCCAGATCGGTCGAGCCCTCAGCAGTTTCTAGTGAA 300

DB 241 CTCGGGGCAAGAACAGATGTCCTCCAGATCGGTCGAGCCCTCAGCAGTTTCTAGTGAA 300

QY 301 TCATCAGATGTTCCAGGGTCCCGCAAGGACCTGAAATGACCTGTACCTTATTGAA 360

DB 301 TCATCAGATGTTCCAGGGTCCCGCAAGGACCTGAAATGACCTGTACCTTATTGAA 360

QY 361 TAAACCAATCAGTTCGCTTCTGCTCTGCTCGCGGCTTCCGCTCTCCAGCTCAATAA 420

DB 361 TAAACCAATCAGTTCGCTTCTGCTCTGCTCGCGGCTTCCGCTCTCCAGCTCAATAA 420

QY 421 AGAGCCCAACCCCTCACTCGGCGCGCAGTCTTCGATAGATGCGTCCCGGGTAC 480

DB 421 AGAGCCCAACCCCTCACTCGGCGCGCAGTCTTCGATAGATGCGTCCCGGGTAC 480

QY 481 CCGTATTCCTCAATAAGGCTCTGCTGTTGCTGATCCGATCGTGTCTCGCTGTTCTTG 540

DB 481 CCGTATTCCTCAATAAGGCTCTGCTGTTGCTGATCCGATCGTGTCTCGCTGTTCTTG 540

QY 541 GGAGGGTCTCTCTGAGTGATTGACTACCCAGCAGGGGGTCTTTTATTTGGGGGCTCGT 600

DB 541 GGAGGGTCTCTCTGAGTGATTGACTACCCAGCAGGGGGTCTTTTATTTGGGGGCTCGT 600

QY 601 CCGGATTTGGAGACCCCTGCGCCAGGGAACACCGACCCAGCCGAGGAGTAACTGGCC 660

DB 601 CCGGATTTGGAGACCCCTGCGCCAGGGAACACCGACCCAGCCGAGGAGTAACTGGCC 660

QY 661 AGCAACTTATCTGTGTCGTGCGATTGTCAGTGTCTATGTTGATGTTATGCGCCCTGCG 720

DB 661 AGCAACTTATCTGTGTCGTGCGATTGTCAGTGTCTATGTTGATGTTATGCGCCCTGCG 720

QY 721 TCTGTACTAGTTAGTAACTAGCTCTGATCTGGCGGACCCGCTGGTGGAACTGACGAGTT 780

DB 721 TCTGTACTAGTTAGTAACTAGCTCTGATCTGGCGGACCCGCTGGTGGAACTGACGAGTT 780

QY 781 CTGAACACCCGGCGCAACCCCTGGGAGACGCTCCAGGGACTTTGGGGCCCGTTTTGTGG 840

DB 781 CTGAACACCCGGCGCAACCCCTGGGAGACGCTCCAGGGACTTTGGGGCCCGTTTTGTGG 840

QY 841 CCGACCTGAGGAGGAGTCGATGTGGAATCCGACCCCGTCAGGATATGTGGTCTTGGT 900

DB 841 CCGACCTGAGGAGGAGTCGATGTGGAATCCGACCCCGTCAGGATATGTGGTCTTGGT 900

QY 901 AGGAGACGAGAACCTAAAAAGTTCCCGCTCCGCTCGAATTTTGTCTTCGGTTGGAA 960

DB 901 AGGAGACGAGAACCTAAAAAGTTCCCGCTCCGCTCGAATTTTGTCTTCGGTTGGAA 960

QY 961 CCGAGCGCGCTTGTCTGTGTCGACCAAGCTTGGGCTGCGAGTCTAGAGGA 1020

DB 961 CCGAGCGCGCTTGTCTGTGTCGACCAAGCTTGGGCTGCGAGTCTAGAGGA 1020

QY 1021 TCAATTCGCGACGAGTAATTCGCTGCTGCTTTTAGGACATATGAAGTATGCACAGT 1080

DB 1021 TCAATTCGCGACGAGTAATTCGCTGCTGCTTTTAGGACATATGAAGTATGCACAGT 1080

QY 1081 GGGATGACTTTTCTGTATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTCG 1140

DB 1081 GGGATGACTTTTCTGTATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTCG 1140

QY 1141 ATGCTCGCTCAGTGAAGCTTTGCTTCTCCGCTTCTTAAATGGCCCACTCTTCAAG 1200

DB 1141 ATGCTCGCTCAGTGAAGCTTTGCTTCTCCGCTTCTTAAATGGCCCACTCTTCAAG 1200

QY 1201 AGAGGATGAAGTCTTATAAACTGACCTGATCACTTTTATCTCATTTGTTGTAGTTC 1260

DB 1201 AGAGGATGAAGTCTTATAAACTGACCTGATCACTTTTATCTCATTTGTTGTAGTTC 1260

QY 1261 TCGTGCCCATCATTTGGCATATAGTCAGCTCAGCTCTGAAATGGGAAACGAAGAAATGCA 1320

DB 1261 TCGTGCCCATCATTTGGCATATAGTCAGCTCAGCTCTGAAATGGGAAACGAAGAAATGCA 1320

QY 1321 CGGTTGGCTCAGTTAATGCGAGATATATCTCAAGTCCGGAAGGCAAGAAATGGCAGTG 1380

DB 1321 CGGTTGGCTCAGTTAATGCGAGATATATCTCAAGTCCGGAAGGCAAGAAATGGCAGTG 1380

QY 1381 AAGATGAATGAGATTCGAGAGAGCTGTGATGGAACGATGAGCAACATGGAAGAGCAAA 1440

DB 1381 AAGATGAATGAGATTCGAGAGAGCTGTGATGGAACGATGAGCAACATGGAAGAGCAAA 1440

QY 1441 TCCAGTATCTTTTCCAGATTAATGAAAGCAATCTCTAGATGCTAAGAAATTTCCAAATTTCA 1500

DB 1441 TCCAGTATCTTTTCCAGATTAATGAAAGCAATCTCTAGATGCTAAGAAATTTCCAAATTTCA 1500

QY 1501 GCATAACACTGATCAAGATTTAATGATGTTTCTTTTCCAGCTAAATTTCTTACTTTCTCT 1560

DB 1501 GCATAACACTGATCAAGATTTAATGATGTTTCTTTTCCAGCTAAATTTCTTACTTTCTCT 1560

QY 1561 CCATCCAGGAAACATGAGATATCATAGGGGATATCTCCAAGTCAATAGTAGGCTCTGAACA 1620

DB 1561 CCATCCAGGAAACATGAGATATCATAGGGGATATCTCCAAGTCAATAGTAGGCTCTGAACA 1620

QY 1621 CCACAGTACTGATTTGAGTTTTCAGTTTGAACACACTGAAATGGCAGAGTCCAGAGAAATG 1680

DB 1621 CCACAGTACTGATTTGAGTTTTCAGTTTGAACACACTGAAATGGCAGAGTCCAGAGAAATG 1680

QY 1681 CATTAAACAAACAGAGGAGATGCGTAAATTTAGAGGAGCGTATATACAATGCAATCAGCAG 1740

DB 1681 CATTAAACAAACAGAGGAGATGCGTAAATTTAGAGGAGCGTATATACAATGCAATCAGCAG 1740

QY 1741 AAATTAAGTCTCTAGATGAAACAAAGTATATTTGGAAACAGGAAATAAAGGGGAAATGA 1800

DB 1741 AAATTAAGTCTCTAGATGAAACAAAGTATATTTGGAAACAGGAAATAAAGGGGAAATGA 1800

QY 1801 AACTGTTGAATATATATCACTAATGATCTGAGGCTGAAAGGATTTGGAAACATTTCTCAGACAT 1860

DB 1801 AACTGTTGAATATATATCACTAATGATCTGAGGCTGAAAGGATTTGGAAACATTTCTCAGACAT 1860

QY 1861 TGAATAATATCACTTTACTCAAGGTCGAGAAAGTCTCGCTGACTGGGAAATGGACCA 1920

DB 1861 TGAATAATATCACTTTACTCAAGGTCGAGAAAGTCTCGCTGACTGGGAAATGGACCA 1920

QY 1921 ACGATCTGGGCTCCAAATGACCATCGGGCTGTGAAACAGCAGAGGTGAATTCACAGGCA 1980
DB 1921 ACGATCTGGGCTCCAAATGACCATCGGGCTGTGAAACAGCAGAGGTGAATTCACAGGCA 1980
QY 1981 CCTACATCACAGCGGTAAACAGCCACATCAAAATGAGATCAAAAGAGTCAACCTGTCATGGGA 2040
DB 1981 CCTACATCACAGCGGTAAACAGCCACATCAAAATGAGATCAAAAGAGTCAACCTGTCATGGGA 2040
QY 2041 CACAAAAACACATCAACAAAGAGGACCCAGCCACCTTTGGCTTCCACCGTCAATTTGGAAGT 2100
DB 2041 CACAAAAACACATCAACAAAGAGGACCCAGCCACCTTTGGCTTCCACCGTCAATTTGGAAGT 2100
QY 2101 TTTTCAGAGTCACACATGCTCTTTCAGGGCCAGTGTCTTCATAGACAGGAAATGGGAAGG 2160
DB 2101 TTTTCAGAGTCACACATGCTCTTTCAGGGCCAGTGTCTTCATAGACAGGAAATGGGAAGG 2160
QY 2161 TCCTGAAAGACCATGTGGCTGTCTGCGGTCAAGTGTAAATGACATTTGGTGTGATGACCTGGAAG 2220
DB 2161 TCCTGAAAGACCATGTGGCTGTCTGCGGTCAAGTGTAAATGACATTTGGTGTGATGACCTGGAAG 2220
QY 2221 CTACCCAGGGTCGGCATCAACATCTTCACTCGCCCTCGGCACACAGAGAGGTGAGTGTG 2280
DB 2221 CTACCCAGGGTCGGCATCAACATCTTCACTCGCCCTCGGCACACAGAGAGGTGAGTGTG 2280
QY 2281 ACCAAGGTCTCTTGGACTCCAGGTGAAAAAGGAGATAGAGCCCTCTCTGGACAAAAATGG 2340
DB 2281 ACCAAGGTCTCTTGGACTCCAGGTGAAAAAGGAGATAGAGCCCTCTCTGGACAAAAATGG 2340
QY 2341 TATACCAAGGCTTTCCAGGTCTAAATAGGTACTCCAGGTCTTAAAGGTGATCGGGGGATCT 2400
DB 2341 TATACCAAGGCTTTCCAGGTCTAAATAGGTACTCCAGGTCTTAAAGGTGATCGGGGGATCT 2400
QY 2401 CTGGTTTACCTGGAGTTCCAGGATTTCCAGGACCAATGGGGAGAGCCGGGAAGCCAGGAC 2460
DB 2401 CTGGTTTACCTGGAGTTCCAGGATTTCCAGGACCAATGGGGAGAGCCGGGAAGCCAGGAC 2460
QY 2461 TTAATGGACAAAAAGCCAGAGGAGGAGAAAAAGGAGTGGAAAGCATGCAAAAGACAATCTA 2520
DB 2461 TTAATGGACAAAAAGCCAGAGGAGGAGAAAAAGGAGTGGAAAGCATGCAAAAGACAATCTA 2520
QY 2521 ATACAGTCCGACTGGTGGGAGCGCCCTCACGAGGAGAGTGGAGATTTTTCACG 2580
DB 2521 ATACAGTCCGACTGGTGGGAGCGCCCTCACGAGGAGAGTGGAGATTTTTCACG 2580
QY 2581 AAGGCCAGTGGGTACCGTGTGTGACGACCGCTCGGAACTCGGTGGAGACTGCTGCTCT 2640
DB 2581 AAGGCCAGTGGGTACCGTGTGTGACGACCGCTCGGAACTCGGTGGAGACTGCTGCTCT 2640
QY 2641 GCAGAGCTTGGGATACAAAGGTGTTCAAAGTGTGCAATAAGCGAGCTTATTTTGGAAAAAG 2700
DB 2641 GCAGAGCTTGGGATACAAAGGTGTTCAAAGTGTGCAATAAGCGAGCTTATTTTGGAAAAAG 2700
QY 2701 GTACGGGTCCAAATATGGCTGAATGAATGATTTTGTTCGGGAAAGAGTCAATTCATGAG 2760
DB 2701 GTACGGGTCCAAATATGGCTGAATGAATGATTTTGTTCGGGAAAGAGTCAATTCATGAG 2760
QY 2761 AGTCGAGAAATAGACAGTGGGTGTGAGAGCTGTTCGACAGCAAGAGTCTCGGGGTC 2820
DB 2761 AGTCGAGAAATAGACAGTGGGTGTGAGAGCTGTTCGACAGCAAGAGTCTCGGGGTC 2820
QY 2821 ACTTTGCACTACATAATGATATATTTTCAATTCACATTTTAAACCTGTATAAAGTG 2880
DB 2821 ACTTTGCACTACATAATGATATATTTTCAATTCACATTTTAAACCTGTATAAAGTG 2880
QY 2881 ATTTTTCCTTTCCTTCACTAAATCAGCTTAATTAATTTTAAAGAACTAAGAAATTTT 2940
DB 2881 ATTTTTCCTTTCCTTCACTAAATCAGCTTAATTAATTTTAAAGAACTAAGAAATTTT 2940
QY 2941 ATCCACAGAAAGGAATATTTTAAATACACTGGATATAATATAATAGCTTCATATTT 3000
DB 2941 ATCCACAGAAAGGAATATTTTAAATACACTGGATATAATATAATAGCTTCATATTT 3000
QY 3001 GCTTCAAATACCAGAACCAATTTCAATCTCTAGGTTTTTAAAGTGGCTCGTGGCGGAATTG 3060

DB 3001 GCTTCAAATACCAGAACCAATTTCAATCTCTAGGTTTTTAAAGTGGCTCGTGGCGGAATTG 3060
QY 3061 ATCCCTTCAGGATATAGTAGTTTCGCTTTTGCATAGGAGGGGAAATGTAGTCTTATGC 3120
DB 3061 ATCCCTTCAGGATATAGTAGTTTCGCTTTTGCATAGGAGGGGAAATGTAGTCTTATGC 3120
QY 3121 AATACTCTTGTAGTCTTTGCAACATGTTAAAGTAGTGTAGCAACATGCTTTACAAGAGA 3180
DB 3121 AATACTCTTGTAGTCTTTGCAACATGTTAAAGTAGTGTAGCAACATGCTTTACAAGAGA 3180
QY 3181 GAAAAACACCGTCATGCGGATTTGGTGAAGTAAAGTGTGACGATCGTCTTATTAGG 3240
DB 3181 GAAAAACACCGTCATGCGGATTTGGTGAAGTAAAGTGTGACGATCGTCTTATTAGG 3240
QY 3241 AAGCAACACAGACGGGTCTGACATGGAATTTGGAACCAACCTGAAATTCGCGATTTGAGAGAT 3300
DB 3241 AAGCAACACAGACGGGTCTGACATGGAATTTGGAACCAACCTGAAATTCGCGATTTGAGAGAT 3300
QY 3301 ATTGTATTTAAGTCTGCTAGCTCGATACAGCAACGCCATTTGACCAATTCACCAATTTGGT 3360
DB 3301 ATTGTATTTAAGTCTGCTAGCTCGATACAGCAACGCCATTTGACCAATTCACCAATTTGGT 3360
QY 3361 GTGCACCTCCAGCTTCAACCTGCGCAGAACCTCAGGGCGCAAGGGCTGCTAAAGGAAG 3420
DB 3361 GTGCACCTCCAGCTTCAACCTGCGCAGAACCTCAGGGCGCAAGGGCTGCTAAAGGAAG 3420
QY 3421 CGGAACACGTAGAAAGCAGTCCGACAGAAACGGTGTGACCCCGGATGAAATGTCAAGCTAC 3480
DB 3421 CGGAACACGTAGAAAGCAGTCCGACAGAAACGGTGTGACCCCGGATGAAATGTCAAGCTAC 3480
QY 3481 TGGGCTATCTGGACAAAGGGGAAAACGCAAGCGCAAGAGAAAGCAGTAGTGTTCAGTGGG 3540
DB 3481 TGGGCTATCTGGACAAAGGGGAAAACGCAAGCGCAAGAGAAAGCAGTAGTGTTCAGTGGG 3540
QY 3541 CTTCATAGGCGATAGCTAGACTGGGCGGTTTTATAGGACAGCAAGCAACCGGAATTTGCCA 3600
DB 3541 CTTCATAGGCGATAGCTAGACTGGGCGGTTTTATAGGACAGCAAGCAACCGGAATTTGCCA 3600
QY 3601 GCTGGGGCGGCTCTGTTGTAAGTTGGGAAGCCCTGCAAAAGTAAACTCGATGGCTTTCTTG 3660
DB 3601 GCTGGGGCGGCTCTGTTGTAAGTTGGGAAGCCCTGCAAAAGTAAACTCGATGGCTTTCTTG 3660
QY 3661 CCGCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCG 3720
DB 3661 CCGCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCG 3720
QY 3721 TTTTCGATGATGAAACAGATGGAATTCACGCGAGGTTCTCGGCGGCTTGGGTGGAGAGG 3780
DB 3721 TTTTCGATGATGGAATTCACGCGAGGTTCTCGGCGGCTTGGGTGGAGAGG 3780
QY 3781 CTATTCGGCTATGACTGGGCAACACAGCAATCGGCTGCTGATGCGCGCTGTTCCGG 3840
DB 3781 CTATTCGGCTATGACTGGGCAACACAGCAATCGGCTGCTGATGCGCGCTGTTCCGG 3840
QY 3841 CTGTACGCGAGGGGCGCCGCTTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTGAT 3900
DB 3841 CTGTACGCGAGGGGCGCCGCTTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTGAT 3900
QY 3901 GAATCGCAGGACGAGGCGCGGCTATCGTGGCTGGCCACGACGGGCTTCTTTGGCA 3960
DB 3901 GAATCGCAGGACGAGGCGCGGCTATCGTGGCTGGCCACGACGGGCTTCTTTGGCA 3960
QY 3961 GCTGTGCTCGAGCTGTGCTCACTGAGCGGAAGGACCTGGCTGCTATTGGGCGAAGTCCG 4020
DB 3961 GCTGTGCTCGAGCTGTGCTCACTGAGCGGAAGGACCTGGCTGCTATTGGGCGAAGTCCG 4020
QY 4021 GGGCAGGATCTCTGTGCTCATCTCACCTTGTCTCGCCGAGAAAGTATCATCATGGCTGAT 4080
DB 4021 GGGCAGGATCTCTGTGCTCATCTCACCTTGTCTCGCCGAGAAAGTATCATCATGGCTGAT 4080
QY 4081 GCAATGCGGGCTGCTCATACGTTGATCCGGCTACCTGCGGCAATTCGACCAACCAAGCGGAA 4140

Db 4081 GCAATGCGCGCTGCATACGCTTGATCCGGCTACCTGCCCCATTCGACCACCAAGCGAAA 4140
Qy 4141 CATGCGATCGAGCGGACGCTACTCGGATGAAGCCGGCTCTGTGCGATCAGGATGATCTG 4200
Db 4141 CATGCGATCGAGCGGACGCTACTCGGATGAAGCCGGCTCTGTGCGATCAGGATGATCTG 4200
Qy 4201 GACGAAGAGCATCAGGGGCTCGCGCCAGCGGAACTGTTGCGCAGGCTCAAGCGCGCATG 4260
Db 4201 GACGAAGAGCATCAGGGGCTCGCGCCAGCGGAACTGTTGCGCAGGCTCAAGCGCGCATG 4260
Qy 4261 CCCGACGCGAGGATCTCGTGAGCCATCGGCATGCGCTCTGTCGCGAATATCATGGTG 4320
Db 4261 CCCGACGCGAGGATCTCGTGAGCCATCGGCATGCGCTCTGTCGCGAATATCATGGTG 4320
Qy 4321 GAAATGSCCCCTTTCTGGAATTCATCGACGTGCGCGGCTGGGTGCGGAGCCGCTAT 4380
Db 4321 GAAATGSCCCCTTTCTGGAATTCATCGACGTGCGCGGCTGGGTGCGGAGCCGCTAT 4380
Qy 4381 CAGGACATAGCGTTGGCTACCGCTGATATTGCTGAAGAGCTTGGCGGCAATGGCTGAC 4440
Db 4381 CAGGACATAGCGTTGGCTACCGCTGATATTGCTGAAGAGCTTGGCGGCAATGGCTGAC 4440
Qy 4441 CGCTTCCTCGCTTTACGCTATCGCGCTCCCGATTCGACGCGCATCGCCTTCTATCGC 4500
Db 4441 CGCTTCCTCGCTTTACGCTATCGCGCTCCCGATTCGACGCGCATCGCCTTCTATCGC 4500
Qy 4501 CTTCTTGACGAGTTCTTCTGAGCGGACTCTGGGGTTGCGATAAATAAAGATTTTATTT 4560
Db 4501 CTTCTTGACGAGTTCTTCTGAGCGGACTCTGGGGTTGCGATAAATAAAGATTTTATTT 4560
Qy 4561 AGTCTCCAGAAAAAGGGGGGAATGAAGACCCCACTGTAGTGTGCGCAAGCTAGCTTAA 4620
Db 4561 AGTCTCCAGAAAAAGGGGGGAATGAAGACCCCACTGTAGTGTGCGCAAGCTAGCTTAA 4620
Qy 4621 GTAAAGCCATTTTCAAGGCTGGAATAATACATACTGAGATAGAGAAAGTTTCAGATCA 4680
Db 4621 GTAAAGCCATTTTCAAGGCTGGAATAATACATACTGAGATAGAGAAAGTTTCAGATCA 4680
Qy 4681 AGGTTCAGGAACAGATGGAACAGCTGGAATATGGGCCAAACAGGATATCTGTGGTAAGCAGT 4740
Db 4681 AGGTTCAGGAACAGATGGAACAGCTGGAATATGGGCCAAACAGGATATCTGTGGTAAGCAGT 4740
Qy 4741 TCCCTGCCCGCTCAGGGCCAAAGACAGATGGAACAGCTGGAATATGGGCCAAACAGGAT 4800
Db 4741 TCCCTGCCCGCTCAGGGCCAAAGACAGATGGAACAGCTGGAATATGGGCCAAACAGGAT 4800
Qy 4801 TCTGTGTAAGCAGTTCTGCCCCGGCTCAGGGCCAAAGACAGATGTTCCCGAGATCGG 4860
Db 4801 TCTGTGTAAGCAGTTCTGCCCCGGCTCAGGGCCAAAGACAGATGTTCCCGAGATCGG 4860
Qy 4861 TCCAGCCCTCAGCAGTTCTTAGAGAACCATCAGATGTTCCAGGGTCCCGCAGGACCTG 4920
Db 4861 TCCAGCCCTCAGCAGTTCTTAGAGAACCATCAGATGTTCCAGGGTCCCGCAGGACCTG 4920
Qy 4921 AAATGACCTGTGCTTATTTGAACTAACCAATCAGTTCGCTTCTGTTCTGTCGCG 4980
Db 4921 AAATGACCTGTGCTTATTTGAACTAACCAATCAGTTCGCTTCTGTTCTGTCGCG 4980
Qy 4981 GCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACACCCCTCACTCGGGGGCGCGAGTCT 5040
Db 4981 GCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACACCCCTCACTCGGGGGCGCGAGTCT 5040
Qy 5041 CCGATTGACTAGTCGCGCGGGTACCGGTGATTCGAATAAACCTCTTGCGAGTTGATCC 5100
Db 5041 CCGATTGACTAGTCGCGCGGGTACCGGTGATTCGAATAAACCTCTTGCGAGTTGATCC 5100
Qy 5101 GACTTGTGGTCTCGCTGTTCTTGGAGGGTCTCTCTGAGTGATGACTACCGGTGAGC 5160
Db 5101 GACTTGTGGTCTCGCTGTTCTTGGAGGGTCTCTCTGAGTGATGACTACCGGTGAGC 5160
Qy 5161 GGGGGTCTTTCAATTTGG 5177
Db 5161 GGGGGTCTTTCAATTTGG 5177

RESULT 2
AAV81410

ID AAV81410 standard; DNA; 7699 BP.

XX AAV81410;

AC AC

XX 17-OCT-2003 (revised)

DT 26-APR-1999 (first entry)

XX Plasmid pLAIN used in retroviral particle delivery.

XX Plasmid pLAIN; retrovirus; gene transfer; mammary gland; milk; lactation;
epithelial cell; trans-somatic animal; amylase; chicken; neo gene;
cyclic; circular; ds.

OS Moloney murine leukemia virus.

OS Gallus sp.

OS Chimeric.

XX CA2224108-A.

XX 05-SEP-1998.

XX 24-FEB-1998; 98CA-02224108.

XX 05-MAR-1997; 97CA-02199212.

XX (MIAC) CANADA MIN AGRICULTURE.

XX Gavora JS, Nguyen TH, Benkel BP, Falconer MM;

XX WPI; 1999-035710/04.

XX Production of a trans-somatic mammal for producing value-added milk - by
transferring the gene into the secretory cells of the mammary gland using
viral derived particles and their packaging cells.

XX Example 1; Page 25-30; 53pp; English.

XX This is the nucleotide sequence of pLAIN, a plasmid designed to carry the
chicken amylase gene (as a marker protein) and various control sequences
for use in retroviral particle delivery to bovine mammary epithelial
cells. The plasmid has the amylase gene and a selectable marker neo gene
both under control of the long terminal repeat promoter of the Moloney
murine leukaemia virus. An internal ribosomal entry site (IRES) provides
a translation initiation site within the transcript, allowing the
downstream gene product to be produced. A method for producing a trans-
somatic mammal for producing an altered composition of milk comprises:
(1) providing a vector containing a DNA sequence encoding a valuable
compound; (2) packaging the vector into a cell line to produce a trans-
infecting particle; (3) preparing a solution comprising the trans-
infecting particle; and (4) delivering the solution into the mammary
gland to allow incorporation of the DNA into the secretory cells of the
mammary gland. The new method is useful for producing milk with valuable
compounds and/or a changed composition, to enhance its properties and/or
marketability. The valuable compound is selected from: an antibody, an
antibiotic, a blood clotting factor, galactosyltransferase, a growth
factor, an oncoprotein, a hormone, a milk protein, a hormone receptor, a
tumour suppressor protein, a vaccine, erythropoietin (claimed), a protein
to aid in weight reduction, a protease inhibitor, an aging inhibitor and
especially is a tissue plasminogen activator which may be extracted or
the milk consumed directly as a therapeutic agent. The new method of
producing trans-somatic animals has an increased trans-infection of
mammary epithelial cells than prior art methods, resulting in an increase
in the concentration and yield of the foreign protein in the milk. The
new method requires only one infusion of the trans-infecting particle,
and one or more genes can be trans-infected. (Updated on 17-OCT-2003 to
standardise OS field)

XX Sequence 7699 BP; 1814 A; 2033 C; 2014 G; 1848 T; 0 U; 0 Other;

Query Match		40.1%;	Score 2077.2;	DB 2;	Length 7699;
Best Local Similarity		64.3%;	Pred. No. 0;		
Matches 3339;		Conservative	0;	Mismatches 1814;	Indels 38; Gaps 14;
Qy	1	TTTGAAGACCCACCCGCTAGCTGGCAAGCTAGCTTAAGTAACGCCACCTTTGCAAGGCAT	60		
Db	175	TTTGAAGACCCACCCGCTAGCTGGCAAGCTAGCTTAAGTAACGCCACCTTTGCAAGGCAT	234		
Qy	61	GGAAAAATACATAAATGAGAAATAGAAAAGTTTCAAGTCAAGGTTCAGGAAACAAAGAAACAGC	120		
Db	235	GGAAAAATACATAAATGAGAAATAGAAAAGTTTCAAGTCAAGGTTCAGGAAACAAAGAAACAGC	294		
Qy	121	TGAATACCAACACAGATATCTGTGGTAAGCGGTTCTGCCCCGCTCAGGCCCAAGAAACA	180		
Db	295	TGAATACCAACACAGATATCTGTGGTAAGCGGTTCTGCCCCGCTCAGGCCCAAGAAACA	354		
Qy	181	GATGAGACAGCTGAGTGAATGGGCCAAACAGAGATATCTGTGGTAAGCGGTTCTGCCCCG	240		
Db	355	GATGAGACAGCTGAGTGAATGGGCCAAACAGAGATATCTGTGGTAAGCGGTTCTGCCCCG	414		
Qy	241	CTCGGGGCCAAGAACAGATGTTCCCAAGATCGGTTCCAGCCCTCAGCAGTTTCTAGTGAA	300		
Db	415	CTCGGGGCCAAGAACAGATGTTCCCAAGATCGGTTCCAGCCCTCAGCAGTTTCTAGTGAA	474		
Qy	301	TCATCAGATGTTTCCAGGTCGCCCAAGGACCTGAAAATGACCTGTACCTTATTTGAAC	360		
Db	475	TCATCAGATGTTTCCAGGTCGCCCAAGGACCTGAAAATGACCTGTACCTTATTTGAAC	534		
Qy	361	TAAACAATCAGTTCGCTTCTCGCTTCTGTTCCGCGCTTCCGCTCCTCGAGCTCAATAAA	420		
Db	535	TAAACAATCAGTTCGCTTCTCGCTTCTGTTCCGCGCTTCCGCTCCTCGAGCTCAATAAA	594		
Qy	421	AGAGCCCAACACCCCTCACTCGGGCGGCCAGTCTTCCGATAGACTTGGCTGCCCGGGTAC	480		
Db	595	AGAGCCCAACACCCCTCACTCGGGCGGCCAGTCTTCCGATAGACTTGGCTGCCCGGGTAC	654		
Qy	481	CCGTATTTCCCAATAAAGCCTTCTGCTGTTTGCATCCGAATCGTGGTCTGCTGTTCTG	540		
Db	655	CCGTATTTCCCAATAAAGCCTTCTGCTGTTTGCATCCGAATCGTGGTCTGCTGTTCTG	714		
Qy	541	GGAGGGTCTCTCTGAGTGATGACTACCAACGAGCGGGGCTTTCATTTGGGGGCTCGT	600		
Db	715	GGAGGGTCTCTCTGAGTGATGACTACCAACGAGCGGGGCTTTCATTTGGGGGCTCGT	774		
Qy	601	CCGGGATTTGGAGACCCCTGCCCCAGGACACACCCACCCAGCGGAGGTAAAGCTGGCC	660		
Db	775	CCGGGATTTGGAGACCCCTGCCCCAGGACACACCCACCCAGCGGAGGTAAAGCTGGCC	834		
Qy	661	AGCAACTTATCTGTCTGTCCGATTTGTCTAGTGTCTATGTTTGAATGTTATGCGCCTGCG	720		
Db	835	AGCAACTTATCTGTCTGTCCGATTTGTCTAGTGTCTATGTTTGAATGTTATGCGCCTGCG	894		
Qy	721	TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGCGGACCCGCTGGTGAACCTGACGAGTT	780		
Db	895	TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGCGGACCCGCTGGTGAACCTGACGAGTT	954		
Qy	781	CTGAACACCCCGCGCAACCCCTGGGAGACGTCGCCAGGACCTTTGGGGCGGTTTCTG	840		
Db	955	CTGAACACCCCGCGCAACCCCTGGGAGACGTCGCCAGGACCTTTGGGGCGGTTTCTG	1014		
Qy	841	CCGACCTGAGGAAGGAGTGCATGTGGAATCCGACCCCGCTCAGGATATGTTGTTCTGTT	900		
Db	1015	CCGACCTGAGGAAGGAGTGCATGTGGAATCCGACCCCGCTCAGGATATGTTGTTCTGTT	1074		
Qy	901	AGGAGACGAGAACCTTAAACAGTTTCCGCGCTCCGCTCTGAATTTTCTGCTTTCGGTTTGAA	960		
Db	1075	AGGAGACGAGAACCTTAAACAGTTTCCGCGCTCCGCTCTGAATTTTCTGCTTTCGGTTTGAA	1134		
Qy	961	CCGAAGCGCGCTCTGTTCTGTCTGTCAGCCCAAGCTTGGGCTGACAGTTCGACTTAGAGGA	1020		
Db	1135	CCGAAGCGCGCTCTGTTCTGTCTGTCAGCCCAAGCTTGGGCTGACAGTTCGACTTAGAGGA	1194		
Qy	1021	TCAATTCGGCACGAGTAAATCGGTCTGCTGCTCTTTAGGACATATGAAGTATGGCACAGT	1080		

Db	1195	GACTGTCTTTCTGTATTTGTTCTGAAATTAAGGCCAGACTGTTACCACCTCCCTTAAGTTT	1254		
Qy	1081	GGGATGACTTTTCTGTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTCCG	1140		
Db	1255	GACCTTAGGTCACTTGGAAAGATGTCGAGCGGATCGCTCACACCCAGTCGGTAGATGTCAA	1314		
Qy	1141	ATGCTCGCTCAGTGACAGCTTTGCTTCTCCCCATCTAAATAATGGGCCCAACTCTTCAAG	1200		
Db	1315	GAAGAGACGTTGGGTTACCTTCTGCTCTGCAGATGCCCAACCTTTAAAGCTCGATGGCC	1374		
Qy	1201	AGAGGATGAAGTCTTATAAAACTGCATGATCACCTTTTATCTCTATTGTGTTTGTAGTTC	1260		
Db	1375	GCAGAGACGGCACCTT-TAACCGAGACCTCATCACCCAGGTTAAGATCAAGGCTCTT-TTC	1431		
Qy	1261	TCGTGCCCATCATTTGGCATAGTGGCAGCTCAGCTCCTGMAATGGGAAACGAAATTTGCA	1320		
Db	1432	ACCTGGCCCGCATGGACACCCAGACAGGTCCCCCTACATCGTGACCTGGGAAGCTTTGGC	1491		
Qy	1321	CGGTTGGCTCAGTTTAATGTCAGATATATCTCAAGTCCGGAAAGCAAAAGAAATGGCAGTG	1380		
Db	1492	TTTGAACCCCTCCCTGGGTCAAGCCCT-----TTGTACACCCCTAAGCCTCCGCTCCT	1546		
Qy	1381	AAGATGAATCAGATTTTCGAGAGCTGTGATGGAAACGATGAGCAACATGGAAGACGAA	1440		
Db	1547	CTTCTCCATCCGCCCCCTCTCTCCCCCTTTGAAACCTCTCGTTCGACCCGCCCTCGATCC	1606		
Qy	1441	TCAGATATCTTTCAGATAATGMAAGCCAAATCTCTAGATGCTTAAGAAATTTCCAAATTTCA	1500		
Db	1607	TCCTTTTATCCAGCCCTCCTCTCTCTTAGGCCCGGAATTCGTTAACTCGACATGGAA	1666		
Qy	1501	GCATAACAACCTGATCAAGATTTTAATGATGTTCTTTTCCAGCTAAATTTCTTACTTCTCT	1560		
Db	1667	GTCTCTTCTCTCTCGCAGCTGTGGGCTTTTGTGGGACACAGTACAAATCCCAACACTCAG	1726		
Qy	1561	CCATCCAGG--AACATGAGAAATATCATAGGGATATCTCCAGTCAATTAGTAGTCTGAA	1618		
Db	1727	GCTGGGAGGACATCTATCGTGCATCTCTTTGAATGGGCTGGGCCGACATTTGCACCTGGAG	1786		
Qy	1619	CACCACAGTACTTGAATTTGAGTTTCAATTTGAAACACTGAAATGCGAGAGTCCCAAGAGAA	1678		
Db	1787	TGCGAACACTATTTAGCTCTTAATGGGTTTGGACGAGTTTCAAGTTTCTCTCCAAATGAA	1846		
Qy	1679	TGCAATTTAAACAAACAGAGGAGATGCGTAAATTAGAGGAGCGTATATACATGATCAGC	1738		
Db	1847	AACATTTGCTACTTAATCCGACAGGCCCTCGTGGGAAAGATACACGCCCTCAGCTAC	1906		
Qy	1739	AGAAATTAAGTCTCTAGATGAAACAAAGTATATTTGGAAACAGGAAATAAAGGGGAAAT	1798		
Db	1907	A-AGATCTGCACCTCGATCGGCCCAATGAAATGAATTCAGAGACATGGTGAACAGATGCAA	1965		
Qy	1799	GAACTGTTTGAATTAATTAATCACTAATGATCTGAGGCTGAAGGATTTGGGACATTTCTCAGAC	1858		
Db	1966	CAATGTTGGAGTTGCTATTTATGTGGATGCTGTTGTTCAATCAGATGTGTGATCTATGGG	2025		
Qy	1859	ATTGAAAAATATCACTTTACTCCAAAGGTGCGAGAAAGTGTCTGCTGACTGGGAAATGAC	1918		
Db	2026	TGGCACGGGCACCACTCAACATGTGGGAGCTATTTCAAACACCGGACTAGAGATTTTCC	2085		
Qy	1919	CAACGATCTGGGCTCCAAATGACCATCTGGGGCTGTGAAACAGACAGAGGTGAATTCACAGG	1978		
Db	2086	CGCTGTCCGCTACTCTGCTGGGATTTCAATGACGGCAAAATGTCACTGCAAGTGGAGA	2145		
Qy	1979	CACCTACATCACGCCCTTAACGCCACATCAATGAGATCAAGAGTCAACAGTCCACCTGATGG	2038		
Db	2146	CATGAAAAATTAATGGGACATGATATCAGGTCGGGATTTGCAAGTTGTCCAGCCTTCTTGA	2205		
Qy	2039	GACAAAAACACCATCAACAGAGGACCCAGGCCACCTTTGGGCTTCA-CGCTCAATTTGA	2097		
Db	2206	CTGCGCTCTGGAGNAGGACTATGTACGCTCAACAAATTCGACGCTACATGATCACCTCAT	2265		
Qy	2098	AGTTTTCAGAGTCCACCACTGTCTTTCACGGGCCAGTCTCTCATAGACAGGAATGGGAAGG	2157		

Db 2266 TGATATGGGTAGCAGGGTTCCGGATCGATGCTGCCAAGCATATGTGGCCAGGGGACAT 2325
Qy 2158 AGGTCTGAAGACCAATGTGGCTGTGGCTCAAGTGTAAATGACATTTGGTGTATGACTGGA 2217
Db 2326 AAGAGCGTTTCTGACAAACCTGCAGATCTAAATACTCAGTGGTTTTTCAGCGAAGCA 2385
Qy 2218 AAGCTACAGGGTC - GGCAATCAACATCTTCACTCGCCTGCGCACACAGAGAGTGTAGTG 2276
Db 2386 ACCCTTTATTACCAAGAGGTAATTGACTTTGGGAGGAGAGCCAAATCAAGGCACTGATG 2445
Qy 2277 AGTGACCAAGTCTCTCGT---GACTCCAGGTGAAAGAGAGATAGAGGCCCTCTCTGAC 2333
Db 2446 CTTTGGGAATGGCGCGTGACAGAAATTCAGATATGTTGTCCTCAAACTGGGGGCGGTGATCCG 2505
Qy 2334 AAAATGGTATACAGGCTTTCCAGGTCTAAATAGGTACTTCCAGGTCTTAAAGGTGATCGGG 2393
Db 2506 GAAGTGGNATCGAGAGAAAGATGGCTACTTAAAGAACTTGGGGAAGGCTGGGGCTTTGT 2565
Qy 2394 GGGATCTCTGGTTTACCTGGAGTTTCGAGGATTCCTCAGGAC--CAATGGGGAAGACCCGGA 2451
Db 2566 GCCTTCTGACAGAGCCCTGTGTGTTGTGGATAACCAAGCAACCAAGCGGGGCGACGGGC 2625
Qy 2452 AGCCAGGACTTAATGGAACAAAGCCNAGAGGAGAAAGAGGAGTGGAAAGCATGCAAA 2511
Db 2626 AGCGGAGCTTCCATTTCTTACTTTCTGGGATGCCAGGCTTTTATAAAATGGCGGTTGGTTT 2685
Qy 2512 GACAATCTAATACAGTCCGACTGTGTGGTGGCGAGCGCCCTCACGAAGGCAGATGGAGA 2571
Db 2686 CATGCTCGCTCATCCGTACGGGTTTCAACGGGTGATGTCAAGTTATCGTTGGCCAAAGATA 2745
Qy 2572 TTTTTCAGAAAGCCAGTGGGTACGGTGTGTGACGACCGCTGGGAACTCGCTGGAGGAC 2631
Db 2746 TTTCCGMAAACCGAGTGGATGTTAACGACTGGGTGGGACCAAGTAACCTGGACGGATC 2805
Qy 2632 T-----GGTGTGTGACGAGCTGGGATACAAAGGTGTTCAAAGTGTGCATATAAGCGAGCT 2687
Db 2806 GACGAAGTCCGTTTCAATCAACGCGAGACATACCTGTGGCAATGACTGGGTCTCGCAACA 2865
Qy 2688 TATTTTGGAAAGGTACGGGTCCAAATATGGCTGAATGAAGTATTTTGTTCGGGAAAGAG 2747
Db 2866 TCGTGGCGCAAAATAGG----ACATGGTTATCTTCCGTAATGTGTAGACGGTCAGC 2921
Qy 2748 TCATCCATTGAAGAGTGCAGAAATTAGACAGTGGGTGTGAGAGCTGTTCGCACAGCA 2807
Db 2922 CTTTCTCAAACTGGTGGGACAAACGGAGCAATCAAGTAGCTTTCCGTGGCGGACAGAG 2981
Qy 2808 GATGTGGGGGTCACTTTGGCACTACATAATGCAATCATATTTTCATTCACATTTTTTAA 2867
Db 2982 GCTTCATTGTCTTTAATAATGATGACTGGTATATGAATGTGCAATTTGCAAACTGGTCTGC 3041
Qy 2868 CTGTTATAAGTGTATTTTTTCTTCTTCTTCACTTAAATCAGCTTAATTAATATTAAGA 2927
Db 3042 CTGCTGGAACTTACTCGCATGTTATTTCTGACAAAAGGAAGGCGAGTGGTGTACTGGAA 3101
Qy 2928 AACTAAGAAATTTTATCCACAGAAAAGGAATATTTAAATAATCACTGGATATAACATATA 2987
Db 3102 AGCAGGTGTACGTTTCTCGATGGAAAGGCCAATTTCCAGATTAGTAAACAGCATGAG 3161
Qy 2988 TAGCTTCATATTTGCTCAAAATACAGAAACATTTCAATTTCTCTAGGTTTTTAAAGTGGC 3047
Db 3162 ATCCATTTGTTGCAATTCACGTTGATGCCAAGTTATAAGCTTTCGAGGATCCACTAGTAAC 3221
Qy 3048 TCGTGGCAATTTGATCCCTCAGGATATAGTATTTTCCGTTTTCATAGGGGGGGGAA 3107
Db 3222 GGCGCCAGTGTGCTGGAAATTCGG-----CTTGTGCAATCTAGGGCGGCGCAATTCGCGC 3276
Qy 3108 TGTAGTCTTATGCAATACTCTTGTAGTCTTGCAACATGGTAACGATGAGTTAGCAACATG 3167
Db 3277 CCTCTCCCCCCCCCTTAAAGTTACTGGCGAGCGGCTTGGAAATAGGCGGCTGTGTG 3336
Qy 3168 CCTTACAGGAGAGAAAGAACCGGTGCATGCCGATTTGGTGGAAAGTAAAGGTGTACGATC 3227
Db 3337 TTTGTCTATATGTGATTTTCCACCATAATTGCCGCTTTTGGCAA-----TGTAGGGGCC 3390

Qy 3228 GTGCTTTATTAGGAAGCAACAGACGGGTCTGACATGGAATTGGACGAACCACTGAATTC 3287
Db 3391 CGGAAACCTGGCCTGTCTTTCTTGACGAGCATTTCTAGGGGTCTTTCCCTCTCGCCAAA 3450
Qy 3288 GCATTTGACAGAGATATTGTAATTAAGTGCCTAGTCTGATACAGCAAAAGCCATTTGACAT 3347
Db 3451 GGAATGCAAGGTCTGTTGAATTCGTTGAAGAGCAGTCTCTCTGGAAGCTTCTTTGAAGA 3510
Qy 3348 TCACCAATTTGGTGTGCACCTCCAAAGCTTCAAGCTCCGCGCGGCAAGCACTCAGGGCGCAAGGG 3407
Db 3511 CAAACAAAGCTGTGTAGCGACCTTTTGACGAGCGGAAACCCCCCACTCTGGCGACAGGTGC 3570
Qy 3408 CTGCTAAAGGAAGCGGAAACAGCTAGAAAGCGAGTCCGACAGAAACGGTGTGACCCCGGAT 3467
Db 3571 CTCTGGGCCCAAAAGCCACGTTATAGATACACCTGCAAAAGGCGGCACAACCCAGTGC 3630
Qy 3468 GAATGTCACTACTGCGGCTATCTGGACA - AGGGAAGAACGCAAGCGCAAGAGAAAGCAGG 3526
Db 3631 CACGTTGTGAGTTGGATAGTTGTGGAAGAGTCAAATGGCTCTCTCAAGCGTAGTCAAC 3690
Qy 3527 TAGCTTGCAGTGGGCTTACATGGCGATAGTACTGGGCGGTTTTTATGGACAGCAAGCG 3586
Db 3691 AAGGGGCTGAAGATGCCAGAAAGTACCCTATTTGTATGGGAATCTGATCTGGGGCCTCG 3750
Qy 3587 AACCGGAATTTGCCAGCTGGGGCGCCCTCTGCTAAGGTTGGGAAGCCCTGCAAAAGTAAACT 3646
Db 3751 GTGCACATGCTTTACATGTGTGTTAGTCGAGGTTTAAAGAGCTTAGGCCCCCGCAAC 3810
Qy 3647 GGAATGGCTTTTCTGCCCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGA 3706
Db 3811 GGGGACGTGTTTTCTTTGAAAACACAGATGATAAGCTTTGCCACAAACCCGGGATAATTC 3870
Qy 3707 CAGGATGAGGATCGTTTCGTCATGATTCGAAAGAGTGAATGGCAGCAGGTTCTCCGGCG 3766
Db 3871 CTGCAGCAATATATGGGATCGGCCATTTGAACAAGATGGAATGGCAGCAGGTTCTCCGGCG 3930
Qy 3767 CTTGGGTGGAGAGGCTATTTCGGCTATGACTGGGCGCAACAGACAACTCGGCTGCTCATG 3826
Db 3931 CTTGGGTGGAGAGGCTATTTCGGCTATGACTTGGGCGCAACAGACAACTCGGCTGCTCATG 3990
Qy 3827 CCGCCGTGTTCCGGCTGTACGCGAGGGGCGCGGTTCTTTTGTCAAGACGACCTGT 3886
Db 3991 CCGCGGTGTTCCGGCTGTACGCGCAGGGGCGCGGTTCTTTTGTCAAGACCGACCTGT 4050
Qy 3887 CCGGTGCCCCGAATGAACTGCAAGACGAGGCGCGGCTATCGTGGCTGGCCACGACGG 3946
Db 4051 CCGGTGCCCCGAATGAACTGCAAGACGAGGCGCGGCTATCGTGGCTGGCCACGACGG 4110
Qy 3947 GCGTTCTTCCGACGCTGTCTCGACGTTGCTCACTGAAGCGGGAAGGAGCTGGCTGCTAT 4006
Db 4111 GCGTTCTTCCGACGCTGTCTCGACGTTGCTCACTGAAGCGGGAAGGAGCTGGCTGCTAT 4170
Qy 4007 TGGCGAAGTCCCGGGCAGGATCTCTGTCTCATCTCACTTGTCTCTGCGCAGAGAAAGTAT 4066
Db 4171 TGGCGAAGTCCCGGGCAGGATCTCTGTCTCATCTCACTTGTCTCTGCGCAGAGAAAGTAT 4230
Qy 4067 CCATCATGCTGTATGCAATCGCGGCGCTGATACGCTTGTATCGGCTACCTGCGCCATTCG 4126
Db 4231 CCATCATGCTGTATGCAATCGCGGCGCTGATACGCTTGTATCGGCTACCTGCGCCATTCG 4290
Qy 4127 ACCACCAAGGAAACATCTGCATCGAGCGACGACGCTACTCGGATGGAAGCGGCTCTTGTG 4186
Db 4291 ACCACCAAGGAAACATCTGCATCGAGCGACGACGCTACTCGGATGGAAGCGGCTCTTGTG 4350
Qy 4187 ATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCGCAACTGTTTCCCGAGGC 4246
Db 4351 ATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCGCAACTGTTTCCCGAGGC 4410
Qy 4247 TCAAGGGCGCATCCCGACGGCGAGGATCTGCTGTGACCCCATGGCGATGCTGCTTGC 4306
Db 4411 TCAAGGGCGCATCCCGACGGCGAGTATCTGCTGTGACCCCATGGCGATGCTGCTTGC 4470

QY 4307 CGAATATCATGCTGGAATATGCGCGCTTTCTGGAATTCATGACTGTGGCGCGCTGGGTG 4366
Db |||||
QY 4471 CGAATATCATGCTGGAATATGCGCGCTTTCTGGAATTCATGACTGTGGCGCGCTGGGTG 4530
Db |||||
QY 4367 TGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGGTGATATTGCTGAAGAGCTTGGCG 4426
Db |||||
QY 4531 TGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGGTGATATTGCTGAAGAGCTTGGCG 4590
Db |||||
QY 4427 GCGAATGGCTGACCGCTTCTCGTGTCTTTACGTTACGCTCCGCTCCCGATTCGAGCGCA 4486
Db |||||
QY 4591 GCGAATGGCTGACCGCTTCTCGTGTCTTTACGTTACGCTCCGCTCCCGATTCGAGCGCA 4650
Db |||||
QY 4487 TCGCTTCTATCGCTTCTTTCAGGAGTCTTCTGAGCGGACTCTGGGTTTCGATAAAT 4546
Db |||||
QY 4651 TCGCTTCTATCGCTTCTTTCAGGAGTCTTCTGAGCGGACTCTGGGTTTCGATAAAT 4710
Db |||||
QY 4547 AAAAGATTTTATTAGTCTCCAGAAAAGGGGGGGAATGAAGACCCACCTGTAGTTTG 4606
Db |||||
QY 4711 AAAAGATTTTATTAGTCTCCAGAAAAGGGGGGGAATGAAGACCCACCTGTAGTTTG 4770
Db |||||
QY 4607 GCAAGCTAGCTTAAGTAACGCCATTTTTCAGAGGATGGAATAATACATACTGAGAATAG 4666
Db |||||
QY 4771 GCAAGCTAGCTTAAGTAACGCCATTTTTCAGAGGATGGAATAATACATACTGAGAATAG 4830
Db |||||
QY 4667 AGAAGTTTCAGATCAAGGTCAGGAACAGATGGAACAGCTGGAATATGGCCAAACAGGATAT 4726
Db |||||
QY 4831 AGAAGTTTCAGATCAAGGTCAGGAACAGATGGAACAGCTGGAATATGGCCAAACAGGATAT 4890
Db |||||
QY 4727 CTGTGGTAAGCAGTTCTGTCGCCCGCTCAGGGCCAGAACAGATGGAACAGCTCAATATG 4786
Db |||||
QY 4891 CTGTGGTAAGCAGTTCTGTCGCCCGCTCAGGGCCAGAACAGATGGAACAGCTCAATATG 4950
Db |||||
QY 4787 GGCCAAACAGGATATCTGTGTAAAGCAGTTCTGTCGCCCGCTCAGGGCCAGAACAGATG 4846
Db |||||
QY 4951 GGCCAAACAGGATATCTGTGTAAAGCAGTTCTGTCGCCCGCTCAGGGCCAGAACAGATG 5010
Db |||||
QY 4847 GTCCCCAGATCGGTCAGCCCTCAGCAGTTCTTAGAGAACCATCAGATGTTCCAGGTT 4906
Db |||||
QY 5011 GTCCCCAGATCGGTCAGCCCTCAGCAGTTCTTAGAGAACCATCAGATGTTCCAGGTT 5070
Db |||||
QY 4907 GCCCAAGGACCTCAATATGACCTGTGCTTATTGAACTAACCAATCAGTTCCCTCTCTC 4966
Db |||||
QY 5071 GCCCAAGGACCTCAATATGACCTGTGCTTATTGAACTAACCAATCAGTTCCCTCTCTC 5130
Db |||||
QY 4967 GCTTCTGTTCGCGGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTC 5026
Db |||||
QY 5131 GCTTCTGTTCGCGGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTC 5190
Db |||||
QY 5027 GGGCGGCGAGTCTCCGATTCAGTTCGCGCGGTTACCGGTGATTCCAATAAACCCCTC 5086
Db |||||
QY 5191 GGGCGGCGAGTCTCCGATTCAGTTCGCGCGGTTACCGGTGATTCCAATAAACCCCTC 5250
Db |||||
QY 5087 TTGCAGTTGCATCCGACTGTGGTCTCGCTTCTTGGGAGGCTCTCCTGAGTGATT 5146
Db |||||
QY 5251 TTGCAGTTGCATCCGACTGTGGTCTCGCTTCTTGGGAGGCTCTCCTGAGTGATT 5310
Db |||||
QY 5147 GACTACCGCTCAGCGGGGCTCTTTCATTTGG 5177
Db |||||
QY 5311 GACTACCGCTCAGCGGGGCTCTTTCATTTGG 5341
Db |||||

RESULT 3

ADQ80675
ID ADQ80675 standard; DNA; 7436 BP.
XX
AC ADQ80675;
XX
DT 21-OCT-2004 (first entry)
XX
DE Bcl-XL expression directing retroviral vector, pPB-bcl-XL-Sn.
XX Survival; neuron; tyrosine hydroxylase; tyrosine 3-monooxygenase; TH;
KW anti-apoptotic; Bcl-XL; neurodegenerative disorder; neuroprotective;

KW nootropic; antiparkinsonian; transplantation; drug screening;
KW gene profiling; CNS disorder; neurodegenerative disease; pPB-bcl-XL-Sn;
KW ds.
XX Unidentified.
XX WO2004062554-A2.
XX
XX 29-JUL-2004.
XX
XX 07-JAN-2004; 2004WO-DK000008.
XX
XX 08-JAN-2003; 2003US-0438719P.
PR 11-APR-2003; 2003DK-00000581.
PR 22-APR-2003; 2003US-0464546P.
XX (NSGE-) NSGENE AS.
XX Martinez-Serrano A, Liste I, Villa A;
XX WPI; 2004-544027/52.
XX
XX Enhancing the survival of neurons or cells expressing tyrosine
PT hydroxylase (TH) for treating neurodegenerative disorders, comprises
PT contacting neurons or TH expressing cells with Bcl-XL or its functional
XX equivalent.
PS Disclosure; SEQ ID NO 6; 108pp; English.

XX The invention relates to a novel method for enhancing the survival of
CC neurons and/or of cells expressing tyrosine hydroxylase (EC 1.14.16.2 -
CC Tyrosine 3-monooxygenase) (TH +). The method comprises contacting a
CC population of cells with Bcl-XL or its functional equivalent, where the
CC population of cells is selected from: neurons or cells capable of
CC differentiating into neurons; or TH expressing cells or cells capable of
CC differentiating into TH expressing cells. The invention further
CC comprises: a composition of cells obtainable by the method above; a
CC composition of isolated mammalian cells overexpressing the anti-apoptotic
CC Bcl-XL protein; a neural progenitor cell; a differentiated dopaminergic
CC neuron; an implantable cell culture device comprising: a semi-permeable
CC membrane permitting the diffusion of a biologically active protein
CC through it; and a composition of cells selected from above; a lentiviral
CC vector particle being produced based on a lentiviral transfer vector;
CC enhancing the survival of TH + cells in vivo; a retroviral particle being
CC produced based on a retroviral transfer vector; enhancing the survival of
CC in vivo differentiated dopaminergic neurons; a packaging cell line
CC capable of producing an infective vector particle; a packaging cell line
CC capable of producing an infective vector particle; treatment of a
CC neurological disorder; a fusion protein comprising the Bcl-XL sequence
CC comprising 233 amino acids ADQ80670 or its functional equivalent and a
CC membrane translocation signal; an expression vector comprising a
CC polynucleotide sequence coding for the fusion protein and a promoter
CC sequence capable of directing the expression of the fusion protein in a
CC host cell; a host cell comprising the expression vector; and producing
CC the fusion protein. The compositions of the invention have
CC neuroprotective, neurotropic, and antiparkinsonian activities. The cells
CC are useful for transplantation, drug screening, gene profiling, or for
CC the preparation of a medicament useful for the treatment of a CNS
CC disorder. The CNS disorder is a neurodegenerative disease involving
CC lesioned and traumatic neurons, including traumatic lesions of peripheral
CC nerves, the medulla, the spinal chord, cerebral ischaemic neuronal
CC damage, neuropathy, peripheral neuropathy, Alzheimer's disease,
CC Huntington's disease, Parkinson's disease, amyotrophic lateral sclerosis,
CC or memory impairment connected to dementia. The method is useful for
CC enhancing the survival of neurons and/or of cells expressing tyrosine
CC hydroxylase for the treatment of neurodegenerative disorders. This
CC sequence represents an expression construct for directing the expression
CC of Bcl-XL in transduced cells, used in the method for increasing the
CC survival rate of neurons of the invention.

XX Sequence 7436 BP; 1744 A; 1993 C; 1939 G; 1760 T; 0 U; 0 Other;

Query Match 38.1%; Score 1972.6; DB 13; Length 7436;

QY 2161 TCCTGAAGACCATGTGGCTGCTGGGTCAAGTGTAAATGACATTTGGTGTATGACTGGAAG 2220
Db |||||
QY 2060 GTGGAGGTCTATAT--AAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGG 2117
Db |||||
QY 2221 CTACCAGGGTCGGCATCAACATCTTCACTCGCTCGGCACACACAGAGAGTGTAGTG 2280
Db |||||
QY 2118 CTTATCGAATTAATACGAC-----TCACTATAGGAGACCCCAAGCTGGCTAGCGTTTAA 2173
Db |||||
QY 2281 ACCAAGGTCTCTCGGACTCCAGGTGMAAAGAGAGATAGAGGCCCTCTCGACACAAATGG 2340
Db |||||
QY 2174 CTTAAGCTTGGTACCGAGCTCGGATCCACTAGTCCAGTGTGGTGAATTTCTGACATATC 2233
QY 2341 TATACAGGCTTTCCAGGTCTAATAGGTACTTCCAGGTCTTAAAGGTGATCGGGGGATCT 2400
Db |||||
QY 2234 CAGCACAGTGGCGCGCTCGATATAAATGTCTCAGAGCAACCGGGAGCTGGTGGTGC 2293
QY 2401 CTGGTTTACCTGGAGTTCCAGGATTTCCAGGACCAATGGGG--AAGACCGGGAAGCCAGG 2458
Db |||||
QY 2294 TTTCTCTCTACAGCTCTCCAGAAAGGATACAGCTGGAGTCAAGTTAGCGATGTCGAA 2353
QY 2459 ACTTAATGGCAAAAAGGCGAGAGAGGAGAGAGAGAGTGGAAAGCATGCAAGACAAATC 2518
Db |||||
QY 2354 GAGAACAGGACTGAAGCCCCAGAGAACTGAAACAGAAAGGAGACCCCAAGTGCATC 2413
QY 2519 TAATACAGTCCGACTGGTGGTGGAGCGGCCCTCAAGAGGCGAGGTGGAGATTTTCA 2578
Db |||||
QY 2414 AATGGCAACCCACTCTGGCACCTGGCGATAGCCCGCGTGAATGGAGCCACTGGCCAC 2473
QY 2579 CGAAGGCCAGTGGGTACGGTGTGACGACCGCTGGGAATCGGTGGAGACTGGTGT 2638
Db |||||
QY 2474 AGCAGAGTTTGGATGCGCGGGAGTACTCCCAATGGCAGCAGTGAAGCAAGCCCTGAGA 2533
QY 2639 CTGCAGGAGCTTGGATACAAAGTGTTCAAAGTGTGCATAAGCGAGCTTATTTTGGAAA 2698
Db |||||
QY 2534 GAGGCTGGCATGAGTTGAACTCGGTGACCGAGAGCAT----- 2573
QY 2699 AGGTACGGGTCCAATATGGCTGAATGAAGTATTTTGTTCGGGAAGAGTCAATCATGA 2758
Db |||||
QY 2574 --TCAGTGTACTAATCCAGCTTCATATAACCCAGGACAGTATATCAGAGCTTTGA 2631
QY 2759 AGAGTGAGATATAGACAGTGGGTGTGAGAGCTGTTCGACACAGAGATGCTGGGG 2818
Db |||||
QY 2632 ACAGGTAGTGAATGAATCTTTTCGGGATGGGGTAAACTGGGGTGCATTTGGCGCTCCTC 2691
QY 2819 TCACCTTGGCACTACATAATGCATCATATTTTCAATTTTCAATTTTAACTGTTTAAAG 2878
Db |||||
QY 2692 CTCCTTTGGGGGGGCACTGTGCGTGGAAAGCGTGAACAAGAGATGCAAGGTATTTGGTGAG 2751
QY 2879 TGATTTTTTCTTTGCTTCACTAAATCAAGTCTTAATTAATTAATTAAGAAACTAAGAAAT 2938
Db |||||
QY 2752 TCGGATTTGCAAGTTGGATGGCCACCTACCTGAATGACCAAC--TAGAGCTTTGGATCCAGG 2810
QY 2939 TTATCCAGAGAAAGGAATATTTAAATCTCGGATAAACAATATAAATAGCTTCATAT 2998
Db |||||
QY 2811 AGAACGGCGGTGGGACACTTTTGTGGATCTCTACGGGAAACAATACAGCACCCGAGA--- 2867
QY 2999 TTGCTTCAATACAGAACCAATTTCAATCTCTAGGTTTTTAAGTGGCTCGTCCCGAT 3058
Db |||||
QY 2868 --GCCGGAAGGCCAGAGAGCGTTTCAACCGCT----GGTTCTGACGGGCACTGTTGGC 2922
QY 3059 TGATCCCTCTCAGGATATAGTATTTGCTTTTGTATAGGAGGGGGAAATGATGCTTAT 3118
Db |||||
QY 2923 TGGTGTAGTTCTGCT--GGGCTCACTCTTCAGTCCGAGTGAACAGATTTCTGAGATAT 2979
QY 3119 GCAATATCTTTGTAGTCTTGCACATGTAACGATAGTATAGCAACATGCTTCAAGGA 3178
Db |||||
QY 2980 CCATCACACTGGCGGCGGAATTCGGGTGAGTCTAGAGGGCCAGATCTAGATCGATAC 3039
QY 3179 GAGAAAGACCTGATGCGGATTTGGTGAAGTGGTGAAGTGGTGGTGGCTTATTA 3238
Db |||||
QY 3040 GTCGACATTTGATCCGCTGTGGAAATGTGTGTCAGTTAGGGGTGGAAAGTCCCGAGCTC 3099

QY 3239 GGAAGGCAACAGACGGGTCTGACATGATTTGGACGAACCACTGAATTCGCAATTCAGAG 3298
Db |||||
QY 3100 CCGAGAGGCAGA----- 3112
QY 3299 ATATTGTATTAAAGTGTAGCTGATACAGCAAAACGCCATTTTGACCAATTCACCAATTCG 3358
Db |||||
QY 3113 -----AGTATGCAAGCATGC 3128
QY 3359 GTGTGCACCTCCAAAGCTTTCAGCTGCGCAAGCACTCAGGGCGCAAGGGCTGTCTAAAGGA 3418
Db |||||
QY 3129 ATCTCAATTAGTCAGCAACAGAGTGTGMAAGTCCCGAGGCTCCAGCAGCGCAGAGTA 3188
QY 3419 AGCGGAACACGTAGAAAGCCAGTCCGCGAGAAACCGGTGTGACCCCGGATGAATGTACGT 3478
Db |||||
QY 3189 TGCAAGC-----ATGCATCTCAAT 3209
QY 3479 ACTGGGCTATCTGACAAAGGGAACCAAGCGCAAGAGAAAGACAGGTAGCTTTCAGTG 3538
Db |||||
QY 3210 AGTCAGCAACCATAGTCCCGCCCTAACTCCGCCCATCCCGCCCTTAACCTCCGCCAGTT 3269
QY 3539 GGCTTACATGCGGATAGCTAGACTGGCGGTGTTTATGACAGCAAGCAACCGGAATTCG 3598
Db |||||
QY 3270 CCGCCCATTTCTCGCCCATGGCTGACTAAATTTTTTTTATTTATGACAGAGCCGA----- 3324
QY 3599 CAGCTGGGGCGCTCTGTGTAAGGTGGGAAGCCCTGCAAG--TAAACTGGATGGCTTT 3656
Db |||||
QY 3325 -GGCGGCTCGGCTCTGAGCTATTCAGAAAGTGTGAGAGGCTTTTTTGGAGSCCTAG 3383
QY 3657 CTTGCCGCCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGATGAGG 3716
Db |||||
QY 3384 GCTTTTCAAAAAGCTTGGGCTGCAGGTGAGGCGGATCTGATCAAGAGACAGATGAGG 3443
QY 3717 ATCGTTTCGATGATGAAACAGATGATGATGACGAGGTTCTCGGCCGCTTGGGTGGA 3776
Db |||||
QY 3444 ATCGTTTCGATGATGAAACAGATGATGATGACGAGGTTCTCGGCCGCTTGGGTGGA 3503
QY 3777 GAGCTATTCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3836
Db |||||
QY 3504 GAGGCTATTCGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3563
QY 3837 CCGGCTGTGACGCGAGGGGCGCGGTTCTTTTGTCAAGACCACTGTCTGGTGGCT 3896
Db |||||
QY 3564 CCGGCTGTGACGCGAGGGGCGCGGTTCTTTTGTCAAGACCACTGTCTGGTGGCT 3623
QY 3897 GATGAACTGACGACGAGCGCGGCTATCGTGTGTCGACGACGAGGCTGTCTGGT 3956
Db |||||
QY 3624 GATGAACTGACGACGAGCGCGGCTATCGTGTGTCGACGAGGCTGTCTGGTGGCT 3683
QY 3957 CGCAGCTGTCTCGACGTTGCTCACTGAAGCGGGAAGGAGTGGCTGTCTATTTGGGCGAAGT 4016
Db |||||
QY 3684 CCGAGCTGTCTCGACGTTGCTCACTGAAGCGGGAAGGAGTGGCTGTCTATTTGGGCGAAGT 3743
QY 4017 GCGGGGCGAGATCTCTGTCTATCTGCTGTCTGTCGCGAGAAAGTATCCATCATGGC 4076
Db |||||
QY 3744 GCGGGGCGAGATCTCTGTCTATCTGCTGTCTGCTGTCTGCTGTCTGCTGTCTGCTGTCT 3803
QY 4077 TGATGAACTGCGGCGCTGATAGCTGTGATCGGCTACCTGCGCATTTCAACCAACCAAGC 4136
Db |||||
QY 3804 TGATGAACTGCGGCGCTGATAGCTGTGATCGGCTACCTGCGCATTTCAACCAACCAAGC 3863
QY 4137 GAAACATTCGATCGAGCGACGATCTCTCGGATGGAAGCGGCTGTCTGATCATCAGGATGA 4196
Db |||||
QY 3864 GAAACATTCGATCGAGCGACGATCTCTCGGATGGAAGCGGCTGTCTGATCATCAGGATGA 3923
QY 4197 TCTGAGCAAGAGATCAGGGGCTCGGCGCAGCGGAATCTGTTGCGCAGGCTCAAGGCGCG 4256
Db |||||
QY 3924 TCTGAGCAAGAGATCAGGGGCTCGGCGCAGCGGAATCTGTTGCGCAGGCTCAAGGCGCG 3983
QY 4257 CATGCGGAGCGGAGATCTCTGTCGTGACCCATGCGGATGCGGCTGTCTGCGGATATCAT 4316
Db |||||
QY 3984 CATGCGGAGCGGAGATCTCTGTCGTGACCCATGCGGATGCGGCTGTCTGCGGATATCAT 4043
QY 4317 GGTGGAATAATGGCGGCTTTTCTCGATTCATCGACTGTGGCGGCTGTGGTGTGGCGGACCG 4376

Db 4044 GGTGAAATGGCGCTTTCTGATTCATCGACTGTGGCGGCTGGGTGGCGGACCG 4103
Qy 4377 CTATCAGACATAGGCTGGCTACCCGCGATATGCTGAAGAGCTTGGCGGCGAATGGGC 4436
Db 4104 CTATCAGACATAGGCTGGCTACCCGCGATATGCTGAAGAGCTTGGCGGCGAATGGGC 4163
Qy 4437 TGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTCGACGCGCATCGCTTCTA 4496
Db 4164 TGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTCGACGCGCATCGCTTCTA 4223
Qy 4497 TCGCTTCTTGAACGAGTTCTTCTGAGCGGGACTCTGGGTTTCGATAAATAAAGATTTT 4556
Db 4224 TCGCTTCTTGAACGAGTTCTTCTGAGCGGGACTCTGGGTTTCGATAAATAAAGATTTT 4283
Qy 4557 ATTTAGTCTCCAGAAAGGGGGNATCAAGAGACCCACCTGTAGTTTGGCAAGCTAGC 4616
Db 4284 ATTTAGTCTCCAGAAAGGGGGNATCAAGAGACCCACCTGTAGTTTGGCAAGCTAGC 4343
Qy 4617 TTAAGTAACGCCATTTTCAAGGCAATGAAATAATACATACTGAGATAGAGAAGTTTCAG 4676
Db 4344 TTAAGTAACGCCATTTTCAAGGCAATGAAATAATACATACTGAGATAGAGAAGTTTCAG 4403
Qy 4677 ATCAAGGTCAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAAG 4736
Db 4404 ATCAAGGTCAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAAG 4463
Qy 4737 CAGTTCTTCCCGGCTCAGGCGCCAGAACAGATGGAACAGCTGAATATGGGCCAAACAG 4796
Db 4464 CAGTTCTTCCCGGCTCAGGCGCCAGAACAGATGGAACAGCTGAATATGGGCCAAACAG 4523
Qy 4797 GATATCTGTGTAGCAGTTCTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCCAAGGA 4856
Db 4524 GATATCTGTGTAGCAGTTCTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCCAAGGA 4583
Qy 4857 GCGGTCCAGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCCAAGGA 4916
Db 4584 GCGGTCCAGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCCAAGGA 4643
Qy 4917 CCTGAATGACCTGTGCTTATTTGAACCAATCAGTTCTGCTTCTGCTTCTGTTTC 4976
Db 4644 CCTGAATGACCTGTGCTTATTTGAACCAATCAGTTCTGCTTCTGCTTCTGTTTC 4703
Qy 4977 GCGGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGGGCGCCAG 5036
Db 4704 GCGGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGGGCGCCAG 4763
Qy 5037 TCCTCCGATTGACTGAGTGCCTCGGTTACCGGTGTATCCAAATAACCCCTTTGCGATTGC 5096
Db 4764 TCCTCCGATTGACTGAGTGCCTCGGTTACCGGTGTATCCAAATAACCCCTTTGCGATTGC 4823
Qy 5097 ATCCGACTTGTGCTTCTGCTTCTGAGAGGCTTCTCTCTGAGTGAATGACTACCGGT 5156
Db 4824 ATCCGACTTGTGCTTCTGCTTCTGAGAGGCTTCTCTCTGAGTGAATGACTACCGGT 4883
Qy 5157 CAGCGGGGCTTTCATTTGG 5177
Db 4884 CAGCGGGGCTTTCATTTGG 4904

RESULT 4

ID AAX90485
XX AAX90485 standard; DNA; 7546 BP.
AC AAX90485;
XX
DT 29-SEP-1999 (first entry)
XX
DE Plasmid retroviral vector pLXSU-IR5-N nucleotide sequence.
KW Plasmid retroviral vector; expression system; immunogenic; gene therapy;
KW immune response; immunosuppression; gene delivery; therapeutic;
KW MHC-I autoimmune disease; tumour; 88.

XX Synthetic.
XX OS
XX PN W09936562-A1.
XX PD 22-JUL-1999.
XX PF 13-JAN-1999; 99WO-US000733.
XX PR 14-JAN-1998; 98US-0071409P.
XX (HUMA-) HUMAN GENE THERAPY RES INST.
XX Radosevich TJ, Link CJ;
XX WPI; 1999-468988/39.
XX Expression system containing therapeutic gene and an immunosuppressor
XX gene useful for treating an MHC-I autoimmune disease or killing tumor
XX cells.
XX Disclosure; Page 147-151; 154pp; English.
XX The present invention describes a nucleotide expression system for the
XX introduction of a therapeutic gene comprising: (i) a nucleotide sequence
XX encoding an immune suppression gene; (ii) a promoter; and (iii) a
XX transcription termination signal, where the system is able to inhibit,
XX evade or eliminate a recipient cell immune response to the therapeutic
XX gene when the gene is transformed into a recipient cell. AAX90481 to
XX AAX90484 represent specifically claimed plasmid retroviral vector
XX nucleotide sequences from the present invention. The expression system
XX and vectors containing it can be used for gene therapy, for treating an
XX MHC-I autoimmune disease or for killing tumor cells. The expression
XX system contains an immunosuppressive gene which prevents host rejection
XX of the vector. The present sequence represents a plasmid retroviral
XX vector from the present invention, designated pLXSU-IR5-N
XX
SQ Sequence 7546 BP; 1695 A; 2066 C; 1981 G; 1804 T; 0 U; 0 Other;
Query Match 37.08; Score 1916.6; DB 2; Length 7546;
Best Local Similarity 64.24; Pred. No. 0;
Matches 3348; Conservative 0; Mismatches 1624; Indels 247; Gaps 21;
Qy 1 TTTGAAAGACCCACCGTAGGTGGCAAGCTAGCTTAAGTAACGCCACTTTGCAAGGCAT 60
Db 175 TTTGAAAGACCCACCGTAGGTGGCAAGCTAGCTTAAGTAACGCCACTTTGCAAGGCAT 234
Qy 61 GGAAAAATACATACTAGAGAAATAGAAAAGTTTCAGATCAAGGTCAAGAACAAAGAAACAGC 120
Db 235 GGAAAAATACATACTAGAGAAATAGAAAAGTTTCAGATCAAGGTCAAGAACAAAGAAACAGC 294
Qy 121 TGAATACCAACAGGATATCTGTGGTAAGCGGTTCTGCCCCGGCTCAGGGCCAAAGAAC 180
Db 295 TGAATACCAACAGGATATCTGTGGTAAGCGGTTCTGCCCCGGCTCAGGGCCAAAGAAC 354
Qy 181 GATGAGACAGCTGAGTGCAGGCGCAACAGGATATCTGTGTAGAGAGTTCTCTCCCGG 240
Db 355 GATGAGACAGCTGAGTGCAGGCGCAACAGGATATCTGTGTAGAGAGTTCTCTCCCGG 414
Qy 241 CTCGGGCGCAAGAACAGATGTTCCCAAGATCGCGTCCAGGCCCTCAGCAGTTTCTAGTGAA 300
Db 415 CTCGGGCGCAAGAACAGATGTTCCCAAGATCGCGTCCAGGCCCTCAGCAGTTTCTAGTGAA 474
Qy 301 TCATCAGATGTTTCCAGGGTGCCTCCCAAGGACCTGAAATGACCCCTGTACCTTATTGAA 360
Db 475 TCATCAGATGTTTCCAGGGTGCCTCCCAAGGACCTGAAATGACCCCTGTACCTTATTGAA 534
Qy 361 TAACCAATCAGTTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTT 420
Db 535 TAACCAATCAGTTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTT 594
Qy 421 AGAGCCCAACACCCCTCAGTCCGCGCGCCAGCTTCTCCGAGTAGACTGCGTCCCGGGTAC 480

Db 595 AGAGCCCAACCCCTCACTCGGCGCCAGTCTTCCGATAGACTGCGTCCCGGGTAC 654
Qy 481 CCGTATTCCTCCCAATAAAGCCTCTTGTCTGTTTGGATCCGAATCGTGGTCTCGCTGTCCTTG 540
Db 655 CCGTATTCCTCCCAATAAAGCCTCTTGTCTGTTTGGATCCGAATCGTGGTCTCGCTGTTCTCTTG 714
Qy 541 GGAGGCTCTCTCTGAGTATGATCTACCCACGAGCGGGGCTTTCATTTTGGGGGCTCGT 600
Db 715 GGAGGCTCTCTCTGAGTATGATCTACCCACGAGCGGGGCTTTCATTTTGGGGGCTCGT 774
Qy 601 CCGGGATTTGGAGACCCCTGCGCCAGGAGCACCCAGCCACCCAGCGGAGGTAAGCTGGCC 660
Db 775 CCGGGATTTGGAGACCCCTGCGCCAGGAGCACCCAGCCACCCAGCGGAGGTAAGCTGGCC 834
Qy 661 AGCAACTTATCTGTCTGTCTGCGATGCTAGTGTCTAAGTGTGATGTTATGCGCCTGCG 720
Db 835 AGCAACTTATCTGTCTGTCTGCGATGCTAGTGTCTAAGTGTGATGTTATGCGCCTGCG 894
Qy 721 TCTGTACTAGTTAGTAACTAGCTGTGATCTGCGGACCCGTTGGTGAACCTGACGAGTT 780
Db 895 TCTGTACTAGTTAGTAACTAGCTGTGATCTGCGGACCCGTTGGTGAACCTGACGAGTT 954
Qy 781 CTGAACACCCCGCGCAACCTGGGAGACGTCCTCAGGACCTTTGGGGCCGTTTTTGTGG 840
Db 955 CTGAACACCCCGCGCAACCTGGGAGACGTCCTCAGGACCTTTGGGGCCGTTTTTGTGG 1014
Qy 841 CCGGACCTGAGGAAGGGAGTGCATGTGGAATCCGACCCCGTCAGGATATGTTCTGCT 900
Db 1015 CCGGACCTGAGGAAGGGAGTGCATGTGGAATCCGACCCCGTCAGGATATGTTCTGCT 1074
Qy 901 AGGAGACGAGAACCTTAAACAGTTTCCCGCCTCGCTCTGAATTTTGTCTCGGTTGGAA 960
Db 1075 AGGAGACGAGAACCTTAAACAGTTTCCCGCCTCGCTCTGAATTTTGTCTCGGTTGGAA 1134
Qy 961 CCGAAGCGCGCGTCTGTCTGTCTGCGACGCAAGCTTGGGCTGCGAGTCTAGAGGA 1020
Db 1135 CCGAAGCGCGCGTCTGTCTGTCTGCGACG----- 1162
Qy 1021 TCATTCGGCACGAGTAAATCGGTGCTGCTGCTTTTAGGACATATGAAGTATGACAGT 1080
Db 1163 -----CGCTGCAGACGCTTCTGTGTGTCTCTGTCTGACTGTGT 1202
Qy 1081 GGGATGACTTTCCTGATCAGCAAGAGGACACTGACGCTGTACAGAGTCTGTGAAGTTG 1140
Db 1203 TTCTGTATTTGTCTGAAATTTAGGGCCAGACTGTATACC-----ACTCCCTTAAGTTTG 1255
Qy 1141 ATGCTCGCTCAGTCAGCTTTGTCTCTCTCCCATCTTAAATGGCCCACTCTTCAAG 1200
Db 1256 ACCTTAGTCTCACTGGAAGAAGTGTGAGCGGATCGCTACAACCAAGTGGTATGATGTCAG 1315
Qy 1201 AGAGGATGAAGTCTTATAAATACTGCATGTGATCACCCTTTATCTCATTTGTGTTGTGTC 1260
Db 1316 AAGAGAGCTTGGGTTA-CCTTCTGCTGTGAGATGGCCAACTTTAAGCTCGATGGCC 1374
Qy 1261 TCGTGCCCATCATTTGGCATATGAGGAGTCTAGCTCTTGAATGGGAAACGAAGAATTTGCA 1320
Db 1375 GCGAGACGGCACCTTTAACCAGAGACCTCATCACCCAGGTTAAGATCAAGTCTTTTCAAC 1434
Qy 1321 CGGTTGGCTCAGTTAATGACATATATCTCAAGTCCGAGGCAAGGAATGGCAGTG 1380
Db 1435 TGGCCCGCATGGACA-----CCGACACAGGTCCTTACATCGTGACCTGG 1480
Qy 1381 AAGATGAATGAGATTTTCGAGAGCTGTGATGGAAACGATGACCAACATGGAAGAGCAAA 1440
Db 1481 GAAGCTTGGCTTTTGAACCCCTCCCTGGGTCAGGCCCTTTGTATACCCCTAAGCCTCG 1540
Qy 1441 TCCAGTATCTTTTCAGATAATGAAGCAATCTCTCTAGATGCTAAGAAATTTCCAAATTTCA 1500
Db 1541 CTTCTCTCTCTCCATCGCCCGGTCTCTCTCCCTTGAACCTCTCTGTTGACCCCGCT 1600
Qy 1501 GCATAACAACTGATCAAGATTTAATGATGTTCTTTTCCAGCTAAATCTCTTCTTCT 1560
Db 1601 CGATCCTCCTTTATCCAGCCCTCACTCTCTCTAGCGCGGAAATTCGTTAACTCGAG 1660

Qy 1561 CCATCCAGGAACATGAGAAATATCATAGGGGATATCTCCAAAGTCAATTAGTAGGCTGGAACA 1620
Db 1661 GATCCGGCTGTGGAAATGTGTCTAGTTAGGGTGTGAAAGTCCCCAGGCTCCCCAGCAGG 1720
Qy 1621 CCACAGTACTTGAATTTTCAGTTTCAGTTTGAACACACTGAATGGGAGAGTCCAAAGAGATG 1680
Db 1721 CAGAAGTA-----TGCAAGCATGCAATCTCAATTTAGTCAGCAACCAAGGTGTGAAAGT 1773
Qy 1681 CATTTAAACAAAGAGGAGATGCGTAAATTTAGAGGAGCGTATATACAAATGCATCAGCAG 1740
Db 1774 CCCCAGGCTCCCGACGAGGCAAGATATGCAAGCATGCAATCTCAATTTAGTCAGCAACA 1833
Qy 1741 AATTAAGTCTCTAGATGAAAAACAAGTATATTTTGGAAACAGAAATAAAGGGGAAATGA 1800
Db 1834 TAGTCCCGCCCTAACTCC--GCCCATCCCGCCCTAACTCCCGCCAGTTCGCGCCCATTC 1891
Qy 1801 AACGTGTTGAATAATCACTAATGATCTGAGGCTGAAGGATTTGGGACATTTCTCAGACAT 1860
Db 1892 TCCGCCCATATGCTGACTAATTTTTTTTATTTATGCAAGAGCCGAGGCCGCTCGGCCCTC 1951
Qy 1861 TGAAAAATATCACTTTACTCCAAAGTCCAGAAAGTCTCGCTGACTGGGAAATGGACCA 1920
Db 1952 TGAGCTA-----TTCCAGAGTATGAGGAGGCTTTTGGAGGCT 1993
Qy 1921 ACGATCTGGGCTCCAAATGACCATCGGGGCTGTGAACAGCAGAGGTGAATTCACAGGCA 1980
Db 1994 AGGCTTTTGCAAAAGCTTGGGCTGCAAGCTTGGTACCGAGCTCGATCGATATCTGCGG 2053
Qy 1981 CCTACATCAGACGCGTAAACAGCCACATCAATAGATCAAAAGTCAACCACTGATGGGA 2040
Db 2054 CCGCGTCGAGCGATGAACCTTGTAAATGCTTATTTAGCCCTCTCGGGCCCGCGGCT 2113
Qy 2041 CACAAAAACACCATCAACAGAGGAGCCAGCCACCTTTGGCTTCCACCGTCAATTTGGAAGT 2100
Db 2114 AGTATGCTGAT-----TATCCTTGNCTTTTTCGNATGAACCTCCGCCCTTGTGTG 2164
Qy 2101 TTTCAGAGTCCACCATCTTCTTACGGGCCAGTCTTCAATAGACAGGAATGGGAAGGAG 2160
Db 2165 GAGACGAGCGGTATACCGCTCTGTCCGATGTTTTCGGAGTACCGAGTAGAGTATTCGAG 2224
Qy 2161 TCCTGAAGACCATGTGGTCTGCTGGTCAAGTGTAAATGATGATTTGGTATGACTGGAAG 2220
Db 2225 GCGCGCTCGCTGCTCCGATCGGGCGGTGCACTGGAGGCTCTGTGGACCTCGCGCGGAA 2284
Qy 2221 CTACAGGGTCGGCATCAACATCTTCACTGCGCTCGCACACAGAGAGTGAAGTGTG 2280
Db 2285 CTGTCCGTGCCACGCCGACA-----CCCCGGGTGTACTACAGACGCTGGAGGCT 2336
Qy 2281 ACCAAGTCTCTCTGAGTCCAGGTGAAAAAGGAGATAGAGGCCCTCTCTGGACAAAAATGG 2340
Db 2337 ACGCGATCGAGTGCAGCCGCGGTGGAGGACGCTCCGAAAGCCTCGTCCGAAA----- 2391
Qy 2341 TATACAGGCTTTTCAGGCTTAATAGGTACTCTCAGGTCTTTAAAGGTGATCGGGGGGATCT 2400
Db 2392 -ACGCTACTGGCTCCGGGACTATCGTGTTCGCCAACGACAAAACTCGTGTGTTGTTACT 2450
Qy 2401 CT--GGTTTACCTGGAGTTCGAGGATTTCCAGGACCAATCGGGGAAGCCCGGAGCCAGG 2458
Db 2451 TTTCCCCCTGCCCAATGCCAACTTATTTATGTAGAGTGGAAACCCCGGTGCTCTGTGC 2510
Qy 2459 ACTTAATGGACAAAAAGCCAGAGGAGGAAAAAGGAGTGGAAAGCATGCAAGAGCAATC 2518
Db 2511 CTTGGGTTCCTCTGTGAGCTCGTTAGAGGNCATCGAAACGACTATTTGTTGGAAGATCGCC 2570
Qy 2519 TAATACAGTCCGACTGTGGTGGGAGCGGCCCTTCAGAAAGGAGAGTGGAGATTTTCA 2578
Db 2571 GTCTAATGGGCTACT-----ACGCGCTCACGATTAAGTCCGCGCAGTATACGCTGATGA 2624
Qy 2579 CGAAGCCAGTGGGTTAGGTTGTGACGACCGCTGGGAACTGCGTGGAGACTGCTGCT 2638
Db 2625 TGGTGGCAGTGAATCAAGTGTGTTTGGGGGCTGTATGTGAAAGGTTGGCTGACCGCAT 2684

QY 2639 CTCGAGGAGCTTGGGATACAAAGGTGTTCAAAGTGTGCATAAGCGAGCTTATTTTGGAAA 2698
Db TTCCCTGGATGTTTTTCGGACCACTGGTGAATTCAGTGCATCCACTAGTACGCCGCCA 2744
QY 2699 AGGTACGGGTCCAATATATGGCTGAATGAAGATATTTTTCGGGAAAGATCATCAATTGA 2758
Db TTCTGCTGGAATTAATTCGCTGTCTGCGAGGCGCGCTGTGTGGGTGAGTACTCCCTCTC 2804
QY 2759 AGAGTGCAGATTAGACAGTGGGGTGTGAGAGCCCTGTCGACGACGAGAGTCTGGGG 2818
Db AAA-----AGCGGCGATGACTTTCGCGCTAAGAT-----TG 2835
QY 2819 TCACCTTTCGACCTACATAATGCATCATATTTTCATTCACATTTTAACTGTTATAAAG 2878
Db TCAGTTTCCAAAACGAGGAGATTGATATCA-----CCTGGCCCGCGGTGATG 2886
QY 2879 TGAATTTTTTCTTGTGCTTCACTAAATCAGCTTAATTAATTAATTAAGAACTAAGAATT 2938
Db CCTTTGAGGGTGGCGCGCTCCATCTGGTCCAGAAAGACAATCTTTTGTGTCAAGCTTG 2946
QY 2939 TTATCCACAGAAAGGATATTTAAATCACTCGNTAAACATATATAAATAGCTTCATAT 2998
Db AGGTGTGCGAGCTTGAGATCTGGCCATACACTTTGAGTGAC-----AATGACATCCACT 3000
QY 2999 TTGCTTCAAATACACAGAACCAATTTCAACTTCTCTAGGTTTTTAAAGTGGCTCGTGCCGAAT 3058
Db TTGCTTTCTCTCCACAGGTGTCCACTCCAGGTCCAACTGCAGGTGCACTGAGCATGCA 3060
QY 3059 TGATCCCTCAGGATATAGTATGTTTCGCTTTTGTGATAGGAGGGGAAATGATGCTTAT 3118
Db TCTAGGGGGGCAATTCGCCCTCTCCCTCCGCCCTTAACGTTACTGCGCGAAGCG 3120
QY 3119 GCATATCTTGTAGCTTGCACATGTAAGTACGATGAGTTAGCAACATGCTTACAAGGA 3178
Db CTTGGAATAAGCCGGTGTGTTTGTCTATGTGATTTTCCACCAT----- 3168
QY 3179 GAGAAAAGCACCGTGATCGGATGTTGGTGAAGTGAAGTGTGATGCTGCTTATTA 3238
Db -----ATTGCCGTCTTTTGGCNAATGTGAGGGCCGGAACCTGGCCCTGTCTTCTTG 3220
QY 3239 GGAAGGCAACAGACGGGTCTGACATGGATTGGACGAACCACTGAAATTCGCAATTTGCAGAG 3298
Db ACGAGCATCTCTAGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTC 3280
QY 3299 ATATTGTTATTTAAGTCCCTAGCTGATACAGCAAAAGCCATTTGACCATTCACACATTTG 3358
Db GTGAAGGAAGCAGTTCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCACCCCTT 3340
QY 3359 GTGTGCACCTCCAAGCTTTCAGCTGCGCAAGCACTCAGGGCGCAAGGGCTGTCTAAAGGA 3418
Db TGCAGGCAAGCGAACCCCTCCTGCGACAGGTGCTCTGCGGCAAAAGCCACGTGTGA 3400
QY 3419 AGCGGAACACGTAGAAAGCCAGTCCGAGAAACGGTGTGACCCCGGATGAATGTACGCT 3478
Db TAAGATACACCTGCAAAGCGGCACT-----AACCCAGTGCACGTTGTGATGTTGGATAG 3455
QY 3479 ACTGGCTATCTGGCAAGGAAACGCAAGCGCAAGAGAAAGCAGGTAGCTTGCAGTG 3538
Db TTGTGGAAGAGTCAATGGCTCTCTCAAGCG-----TAGTCAACAAGGGGCTGAAGGAT 3511
QY 3539 GGCCTTACATGCGATAGCTAGACTGGGCGGTTTTATGGACGACGAGCGAAGCCGAAATTGC 3598
Db GCCCAGAAGGTACCCCATTTGATGGGAATCTGATCTGGGGCTCGGTGCAATGCTTTAC 3571
QY 3599 CAGCTGGGGCCCTCTCGTTAAGGTGGGAAGCCCTGCAAAAGTAACTGGATGGCTTTCT 3658
Db ATGTGTTTAGTCAGGTTTAAAAAGCTCTAGGCCCTTCTAGSCCTCCCGAA--CCACGGGAGCGTGGTTT 3629
QY 3659 TGGCCCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGAT 3718
Db TCCCTTGAATAACACGATGATAAGCTTGGCCAAACCCCGGGATATTTCTCTGACGCCAATA 3689
QY 3719 CGTTTCGCATGATTGAACAAGATGATGTCAGCGAGGTTCTCCGCGCGCTTGGGTGGAGA 3778

Db TGGGATCGGCAATTGAACAAGATGGAATTGCACGAGGTTCTCCGCGCGCTTGGGTGGAGA 3749
QY 3779 GGCTATTTCCGCTATGACTGGGCAACAAGATTCGGCTGCTCTGATGCCGCGCTGTTC 3838
Db GGCTATTTCCGCTATGACTGGGCAACAAGATTCGGCTGCTCTGATGCCGCGCTGTTC 3809
QY 3839 GGCTGTCAAGCGCAGGGGCGCGCTTCTTTTGTCAAGACCGAGCTGTGCGGTGCCCTGA 3898
Db GGCTGTCAAGCGCAGGGGCGCGCTTCTTTTGTCAAGACCGAGCTGTGCGGTGCCCTGA 3869
QY 3899 ATGAATTCGAGGACGAGCGCGCTATCTGTGGCTTGGCCACGACGCGGCTTCTTTCG 3958
Db ATGAATTCGAGGACGAGCGCGCTATCTGTGGCTTGGCCACGACGCGGCTTCTTTCG 3929
QY 3959 CAGCTGTCTCGAGCTTGTCACTGAAGCGGAAGGATCTGGCTGCTATTTGGGCGAAGTGC 4018
Db CAGCTGTCTCGAGCTTGTCACTGAAGCGGAAGGATCTGGCTGCTATTTGGGCGAAGTGC 3989
QY 4019 CGGGGCAAGGATCTCTGCTCATCTCACCTGCTCTGCGGAGAAAGATTCATCATGCTG 4078
Db CGGGGCAAGGATCTCTGCTCATCTCACCTGCTCTGCGGAGAAAGATTCATCATGCTG 4049
QY 4079 ATGCAATCGCGCGCTGCATACGCTTGTATCGGCTACCTGCCCCATTGCACCAACGCA 4138
Db ATGCAATCGCGCGCTGCATACGCTTGTATCGGCTACCTGCCCCATTGCACCAACGCA 4109
QY 4139 AACATCGCATCGAGCAGCAGCTACTCGGATGGAAGCGGCTCTTGTGATCAGATGATC 4198
Db AACATCGCATCGAGCAGCAGCTACTCGGATGGAAGCGGCTCTTGTGATCAGATGATC 4169
QY 4199 TGGACGAAGAGCATCAGGGGCTCGCGCAGCGCAAGCTGTTGCGCAGGCTCAAGCGCGCA 4258
Db TGGACGAAGAGCATCAGGGGCTCGCGCAGCGCAAGCTGTTGCGCAGGCTCAAGCGCGCA 4229
QY 4259 TGCCCGACGCGAGGATCTCGTGTGACCCATGCGCATGCTGCTTCCGCAATATCATGG 4318
Db TGCCCGACGCGAGGATCTCGTGTGACCCATGCGCATGCTGCTTCCGCAATATCATGG 4289
QY 4319 TGAATAATGCGCGCTTTTCTGGAATTCATCGATGTGCGCGCTGGGTGTGGCGACCGCT 4378
Db TGAATAATGCGCGCTTTTCTGGAATTCATCGATGTGCGCGCTGGGTGTGGCGACCGCT 4349
QY 4379 ATCAGGACATAGCGTGTGCTACCGCTGATATGCTGAAGAGCTTGGCGCGCAATGGGCTG 4438
Db ATCAGGACATAGCGTGTGCTACCGCTGATATGCTGAAGAGCTTGGCGCGCAATGGGCTG 4409
QY 4439 ACCGCTTCTCTGCTGCTTACGCTATCGCGCTCCCGATTCGACGCGATCGCCTTCTATC 4498
Db ACCGCTTCTCTGCTGCTTACGCTATCGCGCTCCCGATTCGACGCGATCGCCTTCTATC 4469
QY 4499 GCCTTCTTGAAGGATTC-----TTC 4518
Db GCCTTCTTGAAGGATTC-----TTC 4529
QY 4519 TGAGCGGATCTCGGGTTCGATAAATAAAGATTTTATTTAGTCTCCAGAAAGGG 4578
Db CGCGCGGATCTCGGGTTCGATAAATAAAGATTTTATTTAGTCTCCAGAAAGGG 4589
QY 4579 GGAATGAAGAAGCCCACTCTGCTAGGTTTGGCAAGCTAGCTTAAAGTAAAGGCAATTTGCAAG 4638
Db GGAATGAAGAAGCCCACTCTGCTAGGTTTGGCAAGCTAGCTTAAAGTAAAGGCAATTTGCAAG 4649
QY 4639 GCATGGAAAAATCATAACTGAGATAGAGAGTTTCAGATCAAGGTCAAGNACAGATGGA 4698
Db GCATGGAAAAATCATAACTGAGATAGAGAGTTTCAGATCAAGGTCAAGNACAGATGGA 4709
QY 4699 ACAGCTGAATATGGCCCAACAGGATATCTGTGTAAGAGTTCCTCCCGCGCTCAGGG 4758
Db ACAGCTGAATATGGCCCAACAGGATATCTGTGTAAGAGTTCCTCCCGCGCTCAGGG 4769
QY 4759 CCAAGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTCC 4818

```

Db 4770 CCAAGACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAAGCAGTTCC 4829
QY 4819 TGCCCCGGCTCAGGCCCAAGAACAGAGTGTCCCAAGATGCGGTCCAGCCCTCAGCAGTTT 4878
Db 4830 TGCCCCGGCTCAGGCCCAAGAACAGAGTGTCCCAAGATGCGGTCCAGCCCTCAGCAGTTT 4889
QY 4879 CTAGAGAACCATCAGATGTTTCCAGGTTGCCCAAGGACCTGAATGACCTGTGCTTAA 4938
Db 4890 CTAGAGAACCATCAGATGTTTCCAGGTTGCCCAAGGACCTGAATGACCTGTGCTTAA 4949
QY 4939 TTTGAACCTAACCAATCAGTTCCTCTCTGCTTCTGCTGCGGGCTTCTGCTCCCGAGCT 4998
Db 4950 TTTGAACCTAACCAATCAGTTCCTCTCTGCTTCTGCTGCGGGCTTCTGCTCCCGAGCT 5009
QY 4999 CAATAAAGAGGCCCAACACCCCTCACTCGGGCGCCAGTCTCTCGATTGACTGAGTCGCC 5058
Db 5010 CAATAAAGAGGCCCAACACCCCTCACTCGGGCGCCAGTCTCTCGATTGACTGAGTCGCC 5069
QY 5059 CGGGTACCGGTATCCCAATAAACCCCTTTCGAGTTGCACTCGACTTGTGCTGCTGT 5118
Db 5070 CGGGTACCGGTATCCCAATAAACCCCTTTCGAGTTGCACTCGACTTGTGCTGCTGT 5129
QY 5119 TCCTTGGAGGGTCTCTCTGCTGATGACTACCCGTCAGCGGGGTCTTTCATTGG 5177
Db 5130 TCCTTGGAGGGTCTCTCTGAGTGATGACTACCCGTCAGCGGGGTCTTTCATTGG 5188

RESULT 5
AAQ74450
ID AAQ74450 standard; DNA; 8316 BP.
XX
AC AAQ74450;
XX
DT 22-MAY-1995 (first entry)
XX
DE myoD retroviral expression vector pLHMDN-NSA.
XX
KW MyoD; muscle; gene expression; myoblast; muscle differentiation;
KW rhabdomyosarcoma cell; nuclear regulatory factor; chemotherapeutic agent;
KW identification; detection; skeletal muscle; ss.
XX
OS Synthetic.
XX

Key Location/Qualifiers
FH 1..1159
LTR /*tag= a
FT 1160..1640
FT /*tag= b
FT /*label= Psi region.
FT 1641..2928
FT /*tag= c
FT /*product= "Histidinol dehydrogenase."
FT 2929..4389
FT /*tag= d
FT /*label= MyoD NSA ApaI fragment
FT /*note= "Comprises the myoD proximal and distal regulatory
FT regions."
FT 4390..5259
FT /*tag= e
FT /*product= "Neomycin phosphotransferase."
FT 5260..5964
FT /*tag= f
XX
XX US5352595-A.
XX
XX 04-OCT-1994.
XX
XX 03-SEP-1991; 91US-00753520.
XX
XX 03-SEP-1991; 91US-00753520.
XX
XX (HUTC-) HUTCHINSON CANCER RES CENT FRED.
XX

```

```

PI Palmer TD, Tapscott SJ, Weintraub HM;
XX WPI; 1994-316186/39.
XX
PT Isolated proximal and distal myoD regulatory regions - and vectors contg.
PT them, for tissue selective gene expression in muscle cells and for
XX positive or negative muscle cell selection.
XX Example 8; Fig 8; 48pp; English.
XX
CC The myoD proximal and distal regulatory regions (See AA074447, AA074448)
CC may be inserted into vectors such as this and used (1) to induce a muscle
CC phenotype in a non-muscle cell, (2) for positive or negative selection of
CC muscle cells (3) for targeted gene expression, specifically in skeletal
CC muscle, (4) for the expression of genes controlling growth of myoblasts
CC or mutant forms of myoD that modulate muscle differentiation, (5) for
CC engineering rhabdomyosarcoma cells to increase sensitivity to
CC chemotherapeutic agents, (6) for the expression of growth factors to
CC stimulate skeletal muscle growth in animals, and (7) to identify nuclear
CC regulatory factors that bind to the regulatory region
XX
SQ Sequence 8316 BP; 1922 A; 2246 C; 2255 G; 1880 T; 0 U; 13 Other;
XX
Query Match 32.1%; Score 1661.6; DB 2; Length 8316;
Best Local Similarity 61.1%; Pred. No. 0;
Matches 3536; Conservative 0; Mismatches 1641; Indels 607; Gaps 29;
QY 1 TTTGAAAAGACCCCAACCGGTAGGTGGCAAGCTTAGCTTAAGTAACGCCACTTTGCAAGGCAT 60
Db 175 TTTGAAAAGACCCCAACCGGTAGGTGGCAAGCTTAGCTTAAGTAACGCCACTTTGCAAGGCAT 234
QY 61 GGAAAAATACATAACTCAGAAATAGAAAAGTTTCAGATCAAGGTTCAGGAAACAAGAAACAGC 120
Db 235 GGAAAAATACATAACTCAGAAATAGAAAAGTTTCAGATCAAGGTTCAGGAAACAAGAAACAGC 294
QY 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTCTGCCCCGGCTCAGGCCCAAGAACAA 180
Db 295 TGAATACCAACAGGATATCTGTGTAAGCGGTTCTGCCCCGGCTCAGGCCCAAGAACAA 354
QY 181 GATCAGACAGCTGAGTGTGATGGCCCAACAGGATATCTGTGTAAGCGGTTCTGCCCCGG 240
Db 355 GATGAGACAGCTGAGTGTGATGGCCCAACAGGATATCTGTGTAAGCGGTTCTGCCCCGG 414
QY 241 CTCGGGGCCCAAGAAACAGATGTCTCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGAA 300
Db 415 CTCGGGGCCCAAGAAACAGATGTCTCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGAA 474
QY 301 TCATCAGATGTTTCCAGGGTGGCCCAAGGACCTGAAATGACCTGTACCTTATTGTAAC 360
Db 475 TCATCAGATGTTTCCAGGGTGGCCCAAGGACCTGAAATGACCTGTACCTTATTGTAAC 534
QY 361 TAACCAATCAGTTCTGCTTCTGCTTCTGCTGCGCGGCTTCCGCTCTCGAGCTCAATAAA 420
Db 535 TAACCAATCAGTTCTGCTTCTGCTTCTGCTGCGCGGCTTCCGCTCTCGAGCTCAATAAA 594
QY 421 AGAGCCCAACACCCCTCAGTCTGCGCGCGCCAGTCTTCCGATAGACTGCGTCCCGGGTAC 480
Db 595 AGAGCCCAACACCCCTCAGTCTGCGCGCGCCAGTCTTCCGATAGACTGCGTCCCGGGTAC 654
QY 481 CCGTATTTCCCAATAAAGCCTTCTGCTGTTGTCATCCGAATCGTGTGTTCTGCTT 540
Db 655 CCGTATTTCCCAATAAAGCCTTCTGCTGTTGTCATCCGAATCGTGTGTTCTGCTT 714
QY 541 GGAGGGTCTCTCTGAGTGATGACTACCAACGAGCGGGGTCTTTTCAATTTGGGGGCTCGT 600
Db 715 GGAGGGTCTCTCTGAGTGATGACTACCAACGAGCGGGGTCTTTTCAATTTGGGGGCTCGT 774
QY 601 CCGGATTTGGAGACCCCTCGCCAGGACCAACACCCAGCCACCAACCGGAGGTAACTGCCC 660
Db 775 CCGGATTTGGAGACCCCTCGCCAGGACCAACACCCAGCCACCAACCGGAGGTAACTGCCC 834
QY 661 AGCAACTTATCTGTGTTCTGCTGCGATGTTCTAGTGTCTATGTTTCAATTTGCGCGCTGCG 720

```

Db 835 AGCAACTTATCTGTGTCTGTCGGATTGTCTAGTGTCTATGTTTGTATGCTTATGCGCTGCG 894
Qy 721 TCTGTACTAGTACTAGTACTAGTCTGTATCTGGCGGACCCGTGGTGGAACTGACGAGTT 780
Db 895 TCTGTACTAGTACTAGTACTAGTCTGTATCTGGCGGACCCGTGGTGGAACTGACGAGTT 954
Qy 781 CTGAACACCCCGCCGCAACCTCTGGGAGACGTCCTCAGGAGACTTTGGGGCCCGTTTTGTGG 840
Db 955 CTGAACACCCCGCCGCAACCTCTGGGAGACGTCCTCAGGAGACTTTGGGGCCCGTTTTGTGG 1014
Qy 841 CCCGACCTGAGGAAGGGAGTGCATGTGGAATCCGACCCCGCTCAGGATATGGTTCTGGT 900
Db 1015 CCCGACCTGAGGAAGGGAGTGCATGTGGAATCCGACCCCGCTCAGGATATGGTTCTGGT 1074
Qy 901 AGGAGACGAGAACCTTAAACAGTTTCCCGCTCCGCTCTGAAATTTTTCGTTTGGAA 960
Db 1075 AGGAGACGAGAACCTTAAACAGTTTCCCGCTCCGCTCTGAAATTTTTCGTTTGGAA 1134
Qy 961 CCGAAGCCCGCGCTCTGTCTGTCTGCTGACGCAAGCTTTGGGCTGCAGGTCGACTTAGAGGA 1020
Db 1135 CCGAAGCCCGCGCTCTGTCTGTCTGCTGACGATCGTTCTGTGTGTCTCTGTCTGACTGTGT 1194
Qy 1021 TCAATTCGGCACGAGTAATCGGTCTGCTGCTCTTTAGGACATATGAGTATGGC----- 1075
Db 1195 TTCTGTATTTGTCTGAAATTTAGGSCCAGACTGTTTACCCTCCCTTAAGTTTGAACCTTAG 1254
Qy 1076 -----ACAGTGGGATGACTTTCTGTATGACGAGAGGACACTGACAGCTGTATGAC 1125
Db 1255 GTCACTGGAAGAGTGTGAGGCGGATCGCTCAACACGAGTCGCTAGTCAAGAGAGAC 1314
Qy 1126 AGTCTGTAAGTTCGATCTCGCTCAGTGACAGCTTTGCTTCTCCCTCCCTCTTAAATG 1185
Db 1315 GTTGGTTTACCTTCTGCTCTGAGAAATGCGCAACCTTTAACTGCGATGGCGCGAGACG 1374
Qy 1186 GCCCAACTCTTCAAGAGAGGATGAAGTCTTATAAACTGCACCTGATCACCCTTTATCTCA 1245
Db 1375 GCACCTTTAAACGAGACCTCATCACCAGGTTAAGTCAAGGTCTTTTCACTGGCCGCG 1434
Qy 1246 TTGTGTTTGTAGTCTCTGTGCGCATCATTTGGCATAGTGGGAGCTCAGCTCTGAAATGG 1305
Db 1435 ATGACACCCAGACCCAGGTCCTTACATCGTGTGACCTGGGAGACCTTGGCTTTTGAACCCC 1494
Qy 1306 AAACGAAAGTTCACAGGTTGCTCAGTTAATGACGATATATCTCAAGTCCGGAAGCA 1365
Db 1495 CTCCTGGGTCAAGCCCTTTGTACACCTAAGCTTCGCTCCCTCTTCTCCATCCGCCC 1554
Qy 1366 AAGGAAATGGCAGTGA-----AGATG 1386
Db 1555 CGTCTCTCCCTTGAACCTCTCTGTTCCGACCCCGCTCGATCTCTCCCTTTTATCCAGCCC 1614
Qy 1387 AAATGAGATTTGAGAGACTGTATGGAAAC-----GCATGAGCAACATGGA 1432
Db 1615 TCACCTCTTCTTAGCGCCGGAATTTGATCCGGACCATGAGCTTCAATACCTCGATTGA 1674
Qy 1433 AAGCAGAAATCCAGTATCTTTCAGATAATGAAGCAATCTCTAGATGCTAAGAAATTTCCA 1492
Db 1675 CTGGAACAGCTGTAGCCCTGAAACAGCAGCGGTGCGTGTGAGCGGTCCGGCGATTTCCG 1734
Qy 1493 AAATTTTCAGCATAAAC-----AACTGATCAAGATTTAATGATGTTCTTTTCCAGCT 1543
Db 1735 CTCTGACAGTATACCCGAGCGGTGAGCGATATCTGGATATGTGAAAAACGCGGGTGA 1794
Qy 1544 AAATTCCTTACTTCTCTCCATCCAGGAACATGAGAAATCATAGGG----- 1589
Db 1795 CGATGCCCTTGGTGAATACAGCGCTTAAATTTGATAAAAAACAGAGTGACGCGCTACGCGT 1854
Qy 1590 -----CATATCTCAAGTCAATAGTAGGTCTGAACACCACTGACTTTGATTTGCA 1639
Db 1855 CACCCCTGAAGAGATTCGCGCGCGCGCTGAGCGGCAATTTAAACAGGCGAT 1914
Qy 1640 GTTCAGTATTTGAAACACTGGAATGGCAGAGTCCA---AGAGAAATGCATTTAAACAAACAGA 1696
Db 1915 GACCGCTGCCCTCAAAAAATTTGAAACGTTTCCATTTCCGCGCAGACGCTTACCGCTGTAGA 1974

Qy 1697 GGAGATCGGTAAATTTAGAGGCGGTATATACAATGCATCAGCAGAAATTAAGTCTCTAGA 1756
Db 1975 TGTGAAACCCGACCGCGGTGCTTCCAGCAGGTTACGCGTCCCGTCTCGTCTGTGCG 2034
Qy 1757 T---GAAAAACAAGTATATTTGGAAACAGGAAATAAAAAGGGGAAATGAACCTGTTGAAATA 1813
Db 2035 TCTGTATATTTCCCGCGCGCTCGCTCGCTCTTCTCAACGCTGTGATGCTGGCGACGCC 2094
Qy 1814 TATCATAATATGATCTGAGGCTGAAGGATTTGGGAAACATTTCTCAGACATTTGAAAAATATCAC 1873
Db 2095 GGCGCGAATTTCCGGGATGCGCAGAGGTGGTTCTGTGCTCGCGCCGCCCATCGCTGATGA 2154
Qy 1874 TTTTACTCCAAGGTGCCAGAAAGTCTCGCTGACTGCGGAAATGGACCAAG----- 1923
Db 2155 AATCCTCTATGCGCGCAACTGTGTGCGGTGACGAAATCTTTAACTGCGCGCGCGCGCA 2214
Qy 1924 -----ATCTGGGTCTCCAAATGACCATCGCGGTGTGAAACAGCAGAGGTGAATTCACAGG 1978
Db 2215 GGCGATTTGCGCTCTGGCTTTCGGCAGGAGTCCGTACCGAAAGTGGATAAATTTTGG 2274
Qy 1979 CACCTACATCAGACCGCTTAACAGCACATCAAAATGAGATCA--AAGAGTCAACCTGAT 2036
Db 2275 CCGCGCAACGCTTTTGAACCGAAGCCAAACGTCAGGTCAAGCCAGGCTCTCGACGGCGC 2334
Qy 2037 GGGACACAAAACACCATC-----AACAAAGAGACCCAGCCCACTTTGGCT 2082
Db 2335 GGCTATCGATATGCGACCGCGGCGCTGTGAAGTACTGCTGATCGCAGCAGCGCGCAAC 2394
Qy 2083 TCACCGTCAATTTGGAAGTTTTTACAGAGTCCACCACTGTCTTTCACCGGGC----- 2130
Db 2395 ACCGGATTTGCTGCTCTGACCTGCTCTCCAGGCTGAGCACGCGCGGATTTCCAGGT 2454
Qy 2131 -----AGTGTCTCATAGACAGAAATGGGAAGAGGTCTCTGAAGACCAT 2173
Db 2455 GATCCTGCTGACGCTGATGCTGACATTTGCCGCAAGTGGCGAGGCGGTAGAACGTCA 2514
Qy 2174 GTGGTCTGCGGTCAAGTGTTAATGACATTTGGTGTGATGACTGGGAAAGCTA--CCAGGGTC 2231
Db 2515 ACTGGCGGAACTCGCGCGCGGACACCGCGCGCAGGCTTGAGCGCAGTCTGTGAT 2574
Qy 2232 GGATCAACATCTTCACTCGCTGCGCACACAGAAAGAGGTGAGTGAAGCAACAGGTCTCT 2291
Db 2575 TGTGACCAAGATTTAGCGCAGTGGCTCGCTCTCTAATCAGTATGGCGCGGAACACTT 2634
Qy 2292 CTGGAATCTCAGGTGAAAAAGGAGATAGAGCCCTCTCTGACAA----- 2335
Db 2635 AATCATCCAGACGCGCAATCGCGCGATTTGGTGGATGCGATTTACCAGCGCAGGCTCGGT 2694
Qy 2336 -----AATGGTATACAGGCTTTCAGGCTTAAATAGGTACTCCAGGTCTTAAAGG 2385
Db 2695 ATTTCTCGCGCATGTTGTCGCGGAAATCCGCGGTGATTAACGTTCCGGAACCAACCATGT 2754
Qy 2386 TGATCGGGGGATCTCTGTTTACCTGGAGTTTCAGGATTTCCAGGATTTCCAGGACCAATGGGAAGA 2445
Db 2755 TTTACCGACCTATGGCTATACTGTACTCTGTTCAGGCTTTGGGTAGGGATTTCCAGAA 2814
Qy 2446 CCGGAAAGCCAGGACTTAAATGGACAAAAGG-----CCAGAAAGGA 2486
Db 2815 ACGGATGACCTTCAGGAACTGTGAAAGCGGGCTTTTCGCTCTGCGATCAACCATTTGA 2874
Qy 2487 GAAAAAGGAGTGGAGCATGCAAGACATCTAATACAGTCCGACTGGTGGGTGGGAGC 2546
Db 2875 AACATTTGGCGCGCGCAACGCTGTGACCGCCCAATAAAATGCGGTGACCCCTGCGCGTAAA 2934
Qy 2547 GGCCCTCACAAGSCAGAGTG-----GAGATTTTTTCCAGAAAGCCAGTGGGGTACGGTG 2600
Db 2935 CGCCCTCAAGAGCAACATGAGCACTGAAACACTCTCAGCGTCCGCGGATCAATTCG 2994
Qy 2601 TGTGACGACCGCTGGGAACTGCGTGGAGGACTG--GTGCTGTGACAGGAGCTTGGGATACAA 2659
Db 2995 TTAACTCGAGGATCGGCCCGCAGCCAGGATGGTATAGTATAGNATAGATTTTATTAGA 3054

QY 601 CCGGATTTGGAGACCCCTGCCAGGACCCAGGACCCACCCAGGAGTTAGCTGGCC 660
Db 775 CCGGATTTGGAGACCCCTGCCAGGACCCAGGACCCACCCAGGAGTTAGCTGGCC 834
QY 661 AGCAACTTATCTGTCTGTCCGATTTCTAGTGTCTATGTTGATGTTATGCGCTCGC 720
Db 835 AGCAACTTATCTGTCTGTCCGATTTCTAGTGTCTATGTTGATGTTATGCGCTCGC 894
QY 721 TCTGTACTAGTTAGTAACTAGTCTGTATCTGTATCTGGCGGACCCGTTGGAACTGACAGTT 780
Db 895 TCTGTACTAGTTAGTAACTAGTCTGTATCTGTATCTGGCGGACCCGTTGGAACTGACAGTT 954
QY 781 CTGACACCCCGCCGACCCCTGGAGAGCTCCAGGACCTTGGGGCCCGTTTGTGG 840
Db 955 CTGACACCCCGCCGACCCCTGGAGAGCTCCAGGACCTTGGGGCCCGTTTGTGG 1014
QY 841 CCGGACCTGAGGAGGAGTGCATGTGAATCCGACCCCGTCAGGATATGTGTTCTGTT 900
Db 1015 CCGGACCTGAGGAGGAGTGCATGTGAATCCGACCCCGTCAGGATATGTGTTCTGTT 1074
QY 901 AGGAGACGAGAACCTTAAACAGTTCCTCCGCTCCGTCGAAATTTTGTCTTTCGGTTTGGAA 960
Db 1075 AGGAGACGAGAACCTTAAACAGTTCCTCCGCTCCGTCGAAATTTTGTCTTTCGGTTTGGAA 1134
QY 961 CCGAGCCCGCGCTTCTGTCTGCCTCAGCGCAAGCTTGGGCTGCAGTCTAGAGGA 1020
Db 1135 CCGAGCCCGCGCTTCTGTCTGCCTCAGCGCAAGCTTGGGCTGCAGTCTAGAGGA 1162
QY 1021 TCAATTCGGCAGGATAATCGTCTCGCTCTTTAGGACATATGAAGTATGGCAGT 1080
Db 1163 -----CGCTCAGCATGTTCTGTGTTGTTCTGTCTGCTGACTGTGT 1202
QY 1081 GGGATGACTTTCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTTCG 1140
Db 1203 TTCTGTAATTTGCTGAAATTAGGSCCAGACTGTTACC-----ACTCCCTTAAGTTTG 1255
QY 1141 ATGCTCGCTCAGTGACAGCTTTGCTTCTCCCATCTTAAATAGGCCCAACTCTTTCAAG 1200
Db 1256 ACCTTAGGTCACGTGAAGAGATGTGAGCGGATCGCTCAACACGAGTCTGATAGTCAAG 1315
QY 1201 AGAGATCAAGCTTATAAAGTCACTGATCACCTTTTATCTATTGTTGTTAGTTC 1260
Db 1316 AAGAGACGTTG-----GGTTACCTTCTGCTGTCAGAAATGG 1351
QY 1261 TCGTGCCCATCATTTGGCATAGTGGCAGCTCAGCTCCTGAATTTGGAAACGGAAGATTGCA 1320
Db 1352 CCAACCTTTAAGTCGGATGCGCGGAGACGGCAGCTTTAACCGAGACTCATCA----- 1406
QY 1321 CGTTTGGCTCAGTTAATGCAGATATATCTCAAGTCCGGAAAGGAAATGGCAGTG 1380
Db 1407 -----CCAGGT 1413
QY 1381 AAGATGAATCAGATTTCCAGNAGCTGTGATGGACGATGAGCAACATGGAAGCAGAA 1440
Db 1414 TAAGATCAAGGTCTTTTCACTGGCCCGCATGGACACCCAGACAGGTCCCTCATCATGT 1473
QY 1441 TCCAGTATCTTTAGATATGAAGCAATCTCCTAGATGCTAAGAAATTTCCAAATTTCA 1500
Db 1474 GACCTGGGAAGCCTTGGCTTTTGACCCCTTCCCTGGGTCAAGCCCTTTGTACACCTAA 1533
QY 1501 GCATAACAACTGATCAAAAGATTTAATGATGTTCTTTTCCAGCTAAATCTTACTTCCCT 1560
Db 1534 GCCTCCGCTCTCTTCTCTCATCCGCCCCGTCTCTCCCTTGAACCTCTCTGTTGAC 1593
QY 1561 CCATCCAGAACATGAGATATCATATGAGGATATCTCAAGTCAATAGTATGCTTGACAA 1620
Db 1594 CCGGCTCGATCTCCCTTTAT-----CCAGCCCTCACTCTCTTCTAGG 1638
QY 1621 CCACAGTACTTGAATTTGAGTTTCAATTTGAAACACTGAATGGCAGAGTCCAAGAGATG 1680
Db 1639 CGCCGGAATTCGTT-----AACT 1656
QY 1681 CATTTAAACAAAGAGAGATGCGTAAATTTAGAGGAGCGTATATACAAATGCATCAGCAG 1740

Db 1657 CGAGATCCGCTGTGGAAATGTGTCTAGTTAGGCTGTGGAAGTCCCGAGCTCCCCAG 1716
QY 1741 AAATTAAGTCTCTAGATGAAAAACAAGTATATTTGGAAACAGAAATAAAAGGGGAAATGA 1800
Db 1717 -----CAGGCAGAGTATGCAAGCATGCATCTCTCAATTAGTCAGCA 1757
QY 1801 AACTGTTGAATATATCAATAATGATCTGAGGCTGAAGGATTTGGAAACATTTCTCAGACAT 1860
Db 1758 ACCAGGTGTGAAAGTCTCCAGGCTCCAGCAGCAGCAGAAATATGCAAAA---GCATGCAT 1814
QY 1861 TGAATAATATCACTTTACTTCAAGTGCAGAAAGTGTCTCGTGACTGGGAAATGGACCA 1920
Db 1815 CTCAATTAGTCAGCAACCATAGTCCCGCCCTTAATCCGCCCATCCCGCCCTTAATCTCCG 1874
QY 1921 ACAGTCTGGGCTCCAAACATGACCATCGGGCTGTGAAACAGCAGAGAGTGAATTTACAGGCA 1980
Db 1875 CCGATTTCCGCCCATTTCTCCGCCCATGCTGACTAATTTTTTTTATTTATGACAGAGCC 1934
QY 1981 CCTATACACAGCCGTAAACAGCCATCAAAATGAGATCAAAAGATCACCACTGCATGGGA 2040
Db 1935 GAGCCGCTCGGCTCTGAGCTATTCAGAAAGTAGTGAGGAGGCT-----TTTTTGG 1987
QY 2041 CAAAAACACCATCAACAGAGAGACCCAGCCCACTTTGGCTTCACTGCTCAATTTGGAAGT 2100
Db 1988 AGGCTTAGGCTTTTGAATAAAGCTTGGGCTGCAAGCTTGTACCGAGCTCGATCGATAT 2047
QY 2101 TTTCCAGAGTCCACCACTGTCTTTCAGCGGCCAGTCTTCATAGACAGGAATGGAAGAGG 2160
Db 2048 CTGGCGCGCTGACCGATGTCTGTGGCCCTCG-----AAATGGCGHACCT 2096
QY 2161 TCCTGAAGACCATGTGGCTGTGCGGTCAAGTGTTAATGACATTTGGTGTGATGACATGGAAG 2220
Db 2097 TCCTGGACACCATCGGCTTGGGCCAGGAGTACGCCGACGTACGCCATGA----- 2148
QY 2221 CTACAGGCTCGGCATCAACATCTTCACTCCCTGCGCACACAGAGAGTGAAGTGAAGT 2280
Db 2149 -----GATCAATAAAGGGGCGGTGAGGA 2172
QY 2281 ACCAAGTCTCTGGACTCCAGGTGAANAAGGAGATAGAGCCCTCTCTGACAAAAATGG 2340
Db 2173 CCGGAGCGCGCAGAACCGCCGTGACGACCCCGAGCGTCCCTGTGCGCTCTCTCCGG 2232
QY 2341 TATACAGGCTTTCCAGGTCTAATAGGTACTTCCAGGTCTTAAAGTGTATGCGGGGGAATCT 2400
Db 2233 GCTGCTGCCGGAATCGCCCCAACGCATCTTGTGTTGTCACATCGAAGAACCGGCGG 2292
QY 2401 CTGTTTACCTGGAGTTCGAGGATTTCCAGGACCAATGGGGAGAACCGGGAGCCAGGAC 2460
Db 2293 GACCGTGACCGACAGTCCCGTAAATCCGGTAAACCGTTGAAATTCAGTGGATCC----- 2346
QY 2461 TTAATGGCAAAAAGGCCAGAGGGAGNAAGGGAGTGGAGCATCAAGACAAATCTA 2520
Db 2347 -----ACTA 2350
QY 2521 ATACAGTCCGACTGGTGGTGCGAGCGCCCTCACGAAGCAGAGTGGAGATTTTTCAG 2580
Db 2351 GTAACGCGCGCATGTGCTGG----- 2372
QY 2581 AAGCCCATGTGGGTGACGCTGTGTGACGACCGCTCGGAACTGCGTGGAGGACTGTTGCTGT 2640
Db 2373 -----AATTAATTCGCTGTCTCGAGGCGCGGCTGTTGGGGTGAGTACTCCCTCTCAAAA 2427
QY 2641 GCAGGAGCTTGGGATCAAAAGGTGTTCAAGTGTGCATAGCGAGCTTATTTTCGAAAAAG 2700
Db 2428 GCGGGCATGACTTCTGCGCTAAGATTTGTCAGTTTCCAAAAACAGAGGAGGATTTG----- 2481
QY 2701 GTACGGTCCAAATATGCTGAATGAAGTATTTTGTTCGGGAAAGAGTCACTCCATTGAAG 2760
Db 2482 -----ATATTCACTTCGCCCGC 2498
QY 2761 AGTCAGAAATTTAGACGTGGGTTGTGAGAGCTTGTTCGACGACGAAGAGATCTCGGGGTC 2820

Db 2499 GGTGATGCTTTAGGGTGGCGGCTCCATCTGGTCAGAAAAAGACAATCTTTTGTGTC 2558
Qy 2821 ACTTTGCACTACATAAATGCAATATATTTTCAATTCACATTTTAAACCTGATTAAGTG 2880
Db 2559 AAGCTTAGGTGTGGCAGGCTTGAGATCTGGCCATACACCTTGAGTGACATGACATCCAC 2618
Qy 2881 ATTTTTCCTTTGCTTCACTAAAATCAGCTTAATTAATATATTAAGAACTAAGAAATTTT 2940
Db 2619 TTTGCTTTCTCTCC-----ACAGGT 2639
Qy 2941 ATCCACAGAAAGGAATATTTAAAAATCCTGATATAAATATATAAATAGCTTCATATTT 3000
Db 2640 GTCCATCCCAAGTCCAACTGAGGTGATCGAGCATGCATCCTAGGGCGGCAATTCGCC 2699
Qy 3001 GCTTCAAAATACCAGAACCATTTCAACTTCTCTAGGTTTTAAAGTGGCTCGTGGCGAATTG 3060
Db 2700 CTTCTCCCTCCCGCCCTTAAC-----GTTACTGGCCGA 2734
Qy 3061 ATCCCTCAGGATATAGTAGTTTCGCTTTTGTATAGGAGGGGAAATGTAGTCTTATGC 3120
Db 2735 AGCGCTTGGAAATAGCGCGGTGTGTGTCT-----ATAGTGATTTTCCAC 2784
Qy 3121 AATACTCTTGTAGTCTTGCACATGGTAAACGATGAGTTAGCAACATGCTTACAAAGAGA 3180
Db 2785 CATATTGCCGCTTTTGGCAATGTGAGGGCCCGAAACCTGGCCCTGTCTTCTTGACGAG 2844
Qy 3181 GAAAAGCACCGTGATGCCGATTTGGTGGAAATAGGTGGTACGATCGTGCTTATTAGG 3240
Db 2845 CATTCCTAGGGGTCTTTCCCTCTCGCCAAAGGA-----TGC 2882
Qy 3241 AAGGCAACAGACGGGTCTGACATGATTTGGACCAACCACTGAATTCGCCAATTCAGAGAT 3300
Db 2883 AAGGTCTGTGATGTCTGTAAGGAAGCAGTTCTCTGGAAGCTTCTTGAAGACAAACAA 2942
Qy 3301 ATTTGATTTAAGTGTCTAGCTGATACAGCAAAACGCCAATTTGACCAATTCACATTTGGT 3360
Db 2943 GGTCTGTAGCAACCTTTTGCAGGACAGCGAAACCCCACTTGGCAGCAGGTGCCTCTGCG 3002
Qy 3361 GTGCACCTCCAAGCTTACGCTGCCGCAAGCACTCAGGGCGCAGGGGTGCTAAGGAAG 3420
Db 3003 GCCAAAGCCACGTGTATAAGATACACCTGCAAGGGCGCAC----- 3045
Qy 3421 CGAAACAGGTAGAAACCAAGTCCGAGAAACGGTGTGACCCCGGATGAATGTTCAGCTAC 3480
Db 3046 ---ACCCAGTGCACAGTTGTGAGTTGATGTGTGGAAGTCAATGGCTCTCTC 3102
Qy 3481 TGGGCTATCTGACAAAGGAAACCAAGCGCAAGCAAGAAAGACGAGGTAGCTTGCAGTGGG 3540
Db 3103 AAGCGTAGTCAACAAAGGGGCTGAAGGATGCCAGAAAGGTACCCCAATTTGTATGGGAATCTG 3162
Qy 3541 CTTACATGGGATAGCTAGACTGGCGGTTTATGACAGCAAGCGAACCGGATTTGCCA 3600
Db 3163 AT--CTGGGGCTCGGTGCATGCTTTACATGTTTGTAGTCGAGGTAAAGAGCTCTA 3220
Qy 3601 GCTGGGGCGCTCTGTGTAAGGTTCGGGAAGCCCTGCAAGTAAGCTGATGGCTTTCTTG 3660
Db 3221 GGCCCCCGAACCACGGGGAAGCTGTTTCTTTGAAAACACGATGATAGCTTGCCAC 3280
Qy 3661 CGCCAAAGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATAGGATCG 3720
Db 3281 AACCCCGGAT-----AATTCCTGCAGCCCAATATGGATCG----- 3317
Qy 3721 TTTCCATGATTAACAAGATGGATTGACCGAGGTTCTCCGGCGCTTGGGTGGAGAGG 3780
Db 3318 -----CAATTGAACAAAGATGGATTGACGAGGTTCTCCGGCGCTTGGGTGGAGAGG 3370
Qy 3781 CTTATCGGTATGACTGGGCAACACAGCAATTCGGCTGCTCTGATGCGCGCTGTTCCGG 3840
Db 3371 CTTATTCGGCTATGACTGGGCAACACAGCAATTCGGCTGCTCTGATGCGCGCTGTTCCGG 3430
Qy 3841 CTGTCAAGCGAGGGCGCCGGTCTCTTTGTCAAGACCGACCTGTCCGGTGCCTGAAT 3900
Db 3431 CTGTCAAGCGAGGGCGCCGGTCTCTTTTGTCAAGACCGACCTGTCCGGTGCCTGAAT 3490

Qy 3901 GAACTGCAGGACGAGCGCGCTATCGTGGCTGGCCAGCGGGCGTTCTTTCGCA 3960
Db 3491 GAACTGCAGGACGAGCGCGCTATCGTGGCTGGCCAGCGGGCGTTCTTTCGCA 3550
Qy 3961 GCTGTGCTCGACGCTTGTCACTGAAGCGGAAGGACGTGGCTGCTATTGGGCGAAGTGGCG 4020
Db 3551 GCTGTGCTCGACGCTTGTCACTGAAGCGGAAGGACGTGGCTGCTATTGGGCGAAGTGGCG 3610
Qy 4021 GGGCAGGATCTCCCTGTCATCTCACCTTCTCTCCCGAGGAAGTATCCATCATGGCTGAT 4080
Db 3611 GGGCAGGATCTCCCTGTCATCTCACCTTCTCTCCCGAGGAAGTATCCATCATGGCTGAT 3670
Qy 4081 GCAATGCGGGGCTGTCATACGCTTGATCCGGCTACCTGCCCATTCGACCAACAGCGAAA 4140
Db 3671 GCAATGCGGGGCTGTCATACGCTTGATCCGGCTACCTGCCCATTCGACCAACAGCGAAA 3730
Qy 4141 CATCGATCGAGCGACACGTAFTCGATGGAAGCGGCTTGTGCTGATCGAGATGATCTG 4200
Db 3731 CATCGATCGAGCGACACGTAFTCGATGGAAGCGGCTTGTGCTGATCGAGATGATCTG 3790
Qy 4201 GACGAAGCATCAGGGGCTCGCCGACCGCAACTGTTGCGCAGGCTCAAGCGCGCATG 4260
Db 3791 GACGAAGCATCAGGGGCTCGCCGACCGCAACTGTTGCGCAGGCTCAAGCGCGCATG 3850
Qy 4261 CCCGACGCGAGGATCTCGTGCATGGCGATGCTCTGTTGCCGAATATCATGGTG 4320
Db 3851 CCCGACGCGAGGATCTCGTGCATGGCGATGCTCTGTTGCCGAATATCATGGTG 3910
Qy 4321 GAAATGCGCGCTTTCTGGAATCATCGATCTGCGCGGCTGGGTGTTGGCGGACCGCTAT 4380
Db 3911 GAAATGCGCGCTTTCTGGAATCATCGATCTGCGCGGCTGGGTGTTGGCGGACCGCTAT 3970
Qy 4381 CAGGACATAGGCTTGGCTACCCGATATTTGCTGAAGCTTGGGGCGAATGGCTGAC 4440
Db 3971 CAGGACATAGGCTTGGCTACCCGATATTTGCTGAAGCTTGGCGGGAATGGCTGAC 4030
Qy 4441 CGCTTCTCGTGTCTTACGGTATCGCGCTCCGATTCGACGCGCATCGCTTCTTATCGC 4500
Db 4031 CGCTTCTCGTGTCTTACGGTATCGCGCTCCGATTCGACGCGCATCGCTTCTTATCGC 4090
Qy 4501 CTTCTTGAAGGTTTC-----TTCTG 4520
Db 4091 CTTCTTGAAGGTTTCCTGTCGAGCGGATCTGATCAAGAGACAGGATGAGGATCGTTTCG 4150
Qy 4521 AGCGGACTCTGGGTTGATATAAATAGATTTTATTTAGTCTCCAGRAAAGGGGG 4580
Db 4151 CGCGGACTCTGGGTTGATATAAATAGATTTTATTTAGTCTCCAGRAAAGGGGG 4210
Qy 4581 AATGAAAGACCCCACTGTAGTTTGGCAAGCTAGCTTAAGTAAACGCCATTTTGAAGGC 4640
Db 4211 AATGAAAGACCCCACTGTAGTTTGGCAAGCTAGCTTAAGTAAACGCCATTTTGAAGGC 4270
Qy 4641 ATGAAATAATACATAAATGAGAAATGAGATTCAGATCAAGGTGAGAAACAGATGGAAC 4700
Db 4271 ATGAAATAATACATAAATGAGAAATGAGATTCAGATCAAGGTGAGAAACAGATGGAAC 4330
Qy 4701 AGCTGAATATGGGCAACAGGATATCTGTGGTGAAGATTTCTTCCGCCCGCTCAGGGCC 4760
Db 4331 AGCTGAATATGGGCAACAGGATATCTGTGGTGAAGATTTCTTCCGCCCGCTCAGGGCC 4390
Qy 4761 AAGAAACAGATGGAACAGCTGAATATGGGCAACAGGATATCTGTGGTGAAGATTTCTTCC 4820
Db 4391 AAGAAACAGATGGAACAGCTGAATATGGGCAACAGGATATCTGTGGTGAAGATTTCTTCC 4450
Qy 4821 CCCCGCTCAGGGCCAAAGAACAGATGTTCCAGATGCGGTCCAGCTCAGAGCTTCT 4880
Db 4451 CCCCGCTCAGGGCCAAAGAACAGATGTTCCAGATGCGGTCCAGCTCAGAGCTTCT 4510
Qy 4881 AGAGAACATCAGATGTTTCCAGGGTGGCCCAAGGACCTGAATGACCTGTGCTTATTT 4940
Db 4511 AGAGAACATCAGATGTTTCCAGGGTGGCCCAAGGACCTGAATGACCTGTGCTTATTT 4570

QY 4941 TGAACCTAACCAATCAGTTGGCTTCTCGCTTCTGTTGGCGGCTTCTGCTCCCGAGCTCA 5000
Db |||||
QY 4571 TGAACCTAACCAATCAGTTGGCTTCTCGCTTCTGTTGGCGGCTTCTGCTCCCGAGCTCA 4630
Db |||||
QY 5001 ATAAAGAGAGCCCAACACCCCTCACTCGGGGCGCCAGTCTCCGATTGACTGAGTGGCCCG 5060
Db |||||
QY 4631 ATAAAGAGAGCCCAACACCCCTCACTCGGGGCGCCAGTCTCCGATTGACTGAGTGGCCCG 4690
Db |||||
QY 5061 GGTACCCGTTGATCCAAATAAACCCCTCTTGCAGTTGCATCCGACTTGTGGTCTCGCTGTC 5120
Db |||||
QY 4691 GGTACCCGTTGATCCAAATAAACCCCTCTTGCAGTTGCATCCGACTTGTGGTCTCGCTGTC 4750
Db |||||
QY 5121 CTTGGAGGGTCTCTCTGAGTGATTGACTACCCGTCAGCGGGGTCTTTTCAATTGG 5177
Db |||||
QY 4751 CTTGGAGGGTCTCTCTGAGTGATTGACTACCCGTCAGCGGGGTCTTTTCAATTGG 4807
Db |||||

RESULT 7

AAV33629

ID AAV33629 standard; DNA; 5594 BP.

XX AC

XX AAV33629;

XX DT 29-DEC-1998 (first entry)

XX XX

XX DE GENSA 981, a monomeric DNA sequence produced by the invention.

XX KW

XX KW Class IIS restriction endonuclease recognition site;

XX KW endogenous mouse promoter element; tissue-specific gene expression;

XX KW hormone-specific gene expression; ss;

XX KW developmental-specific gene expression.

XX OS Synthetic.

XX XX

XX PN WO9838326-A1.

XX PD 03-SEP-1998.

XX PF 28-FEB-1998; 98WO-US003918.

XX PR 28-FEB-1997; 97US-0070910P.

XX XX

XX PA (NATU-) NATURE TECHNOLOGY CORP.

XX XX

XX PI Hodgson CP, Zink MA, Xu G;

XX XX

XX DR WPI, 1998-495399/42.

XX PT Method for assembling gene or gene vector - comprises use of primers

XX PT containing class IIS restriction endonuclease recognition sites.

XX XX

XX PS Example 2; Page 107-110; 141pp; English.

XX XX

XX CC The invention provides a novel method for directing self-assembly of a

XX CC gene having three or more fragments in a directionally and spatially

XX CC ordered fashion to produce a gene or a gene vector. The method involves

XX CC usage of primers, containing class IIS restriction endonuclease

XX CC recognition sites, for isolation of these fragments. As described in the

XX CC disclosure, the method may also use a vector for the incorporation and

XX CC screening of endogenous mouse promoter elements for the identification of

XX CC cell specific promoters. In the example given, plasmids pBK-CMV

XX CC (AAV33626), pVLM (AAV33623) and pVLOVHGH-900 (AAV33621) were used as

XX CC templates from which six fragments were amplified. Each of the fragments

XX CC contained different regulatory sequences. The six PCR fragments were

XX CC designed to self-assemble into a retro-vector using the method of the

XX CC invention. The present sequence, designated as GENSA 981, represents the

XX CC monomeric DNA sequence of the six ligated fragments. In general, the

XX CC method is claimed to be useful for isolating and identifying regulatory

XX CC sequences from a cell, including those for enhanced biological activity,

XX CC or tissue-specific, hormone-specific or developmental-specific gene

XX CC expression

XX SQ

XX Sequence 5594 BP; 1232 A; 1519 C; 1518 G; 1325 T; 0 U; 0 Other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Query Match 31.5%; Score 1629.4; DB 2; Length 5594;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1654; Conservative 0; Mismatches 21; Indels 2; Gaps 1;
QY 3371 AAGCTTCACTGCGCCGCAAGCACTCAGGGCCCAAGGGCTGCTAAAGGAAGCGGACACGT 3430
Db |||||
QY 3342 AAGCTTCACTGCGCCGCAAGCACTCAGGGCCCAAGGGCTGCTAAAGGAAGCGGACACGT 3401
Db |||||
QY 3431 AGAAAGCAGTCCCGCAGAAACGGTCTGACCCCGGATGAATGTCACTACTGGGCTATCT 3490
Db |||||
QY 3402 AGAAAGCAGTCCCGCAGAAACGGTCTGACCCCGGATGAATGTCACTACTGGGCTATCT 3461
Db |||||
QY 3491 GGAAAGGAGGAAAAACGAAAGCGAAAGAGAGAGTGTGAGTGGGCTTACATGGC 3550
Db |||||
QY 3462 GGAAAGGAGGAAAAACGAAAGCGAAAGAGAGAGTGTGAGTGGGCTTACATGGC 3521
Db |||||
QY 3551 GATAGCTAGACTGGGCGGTTTTATGGACAGCAAGCGGAACCGGATTTGCCAGCTGGGGCGC 3610
Db |||||
QY 3522 GATAGCTAGACTGGGCGGTTTTATGGACAGCAAGCGGAACCGGATTTGCCAGCTGGGGCGC 3581
Db |||||
QY 3611 CCTCTGGTAAGTTGGGAAGCCCTGCAAGTAAACTGGATGGCTTTCTTGGCGCCCAAGGA 3670
Db |||||
QY 3582 CCTCTGGTAAGTTGGGAAGCCCTGCAAGTAAACTGGATGGCTTTCTTGGCGCCCAAGGA 3641
Db |||||
QY 3671 TCTGATGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCGTTTTCGCATGA 3730
Db |||||
QY 3642 TCTGATGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCGTTTTCGCATGA 3701
Db |||||
QY 3731 TTGAACAAGATGGAATGCAACGAGGTTCTCCGGCGGCTTGGGTTGAGAGGCTATTCGGCT 3790
Db |||||
QY 3702 TTGAACAAGATGGAATGCAACGAGGTTCTCCGGCGGCTTGGGTTGAGAGGCTATTCGGCT 3761
Db |||||
QY 3791 ATGACTGGGCACAACAGACAACTCGGCTGCTGATGCGCGCTGTTCCGGCTGTCCAGGC 3850
Db |||||
QY 3762 ATGACTGGGCACAACAGACAACTCGGCTGCTGATGCGCGCTGTTCCGGCTGTCCAGGC 3821
Db |||||
QY 3851 AGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTGTAATGAATGCAGG 3910
Db |||||
QY 3822 AGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTGTAATGAATGCAGG 3881
Db |||||
QY 3911 ACAGGACAGCGCGGCTATCGTGGCTGGCCAGACGCGGGGTTCTTGGCGACGTGCTCG 3970
Db |||||
QY 3882 ACAGGACAGCGCGGCTATCGTGGCTGGCCAGACGCGGGGTTCTTGGCGACGTGCTCG 3941
Db |||||
QY 3971 ACCTTTGTCACCTGAAGCGGAGAGGACTGGCTGCTATTGGGCGAAGTGCCGGGAGGATC 4030
Db |||||
QY 3942 ACCTTTGTCACCTGAAGCGGAGAGGACTGGCTGCTATTGGGCGAAGTGCCGGGAGGATC 4001
Db |||||
QY 4031 TCCTGTGTCATCTCACCTTGTCTCTCCCGAGAAAGTATCCATCATGCTGATGCAATGCAGC 4090
Db |||||
QY 4002 TCCTGTGTCATCTCACCTTGTCTCTCCCGAGAAAGTATCCATCATGCTGATGCAATGCAGC 4061
Db |||||
QY 4091 GGCTGCATACCTGTGATCCGGCTACCTGCCCATCTTCGACCAACGAGCGGAAACATGCGCATCG 4150
Db |||||
QY 4062 GGCTGCATACCTGTGATCCGGCTACCTGCCCATCTTCGACCAACGAGCGGAAACATGCGCATCG 4121
Db |||||
QY 4151 AGCGGACAGCTACTCCGATGGAAGCGGCTCTTGTGATCAGGATGATCTGGACGAGGAGC 4210
Db |||||
QY 4122 AGCGGACAGCTACTCCGATGGAAGCGGCTCTTGTGATCAGGATGATCTGGACGAGGAGC 4181
Db |||||
QY 4211 ATCAGGGGCTCGCGCCACGAGCCGAACTGTTTCCGACAGGCTCAAGCGCGCATGCCCCACGCGC 4270
Db |||||
QY 4182 ATCAGGGGCTCGCGCCACGAGCCGAACTGTTTCCGACAGGCTCAAGCGCGCATGCCCCACGCGC 4241
Db |||||
QY 4271 AGGATCTCGTGTGACCCATGCGCATGCTCTTCCGGAATATCATGTTGAAATGCGC 4330
Db |||||
QY 4242 AGGATCTCGTGTGACCCATGCGCATGCTCTTCCGGAATATCATGTTGAAATGCGC 4301
Db |||||
QY 4331 GCTTTTCTGGATTCATCGACTGTGGCGGCTGGGGTGTGGCGGACCGCTATCAGGACATAG 4390
Db |||||
QY 4302 GCTTTTCTGGATTCATCGACTGTGGCGGCTGGGGTGTGGCGGACCGCTATCAGGACATAG 4361
Db |||||

QY 61 GGAATAATACATACTGAGATAGAAAAAGTTTCAGATCAAGGTTCAGGAACAAGAAACAGC 120
DB |||||
QY 61 GGAATAATACATACTGAGATAGAAAAAGTTTCAGATCAAGGTTCAGGAACAAGAAACAGC 120
DB |||||
QY 121 TGAATACCAACACAGGATATCTGTGTAAAGCGTTCTCCCGGCTCAGGSCCAAGAAC 180
DB |||||
QY 121 TGAATACCAACACAGGATATCTGTGTAAAGCGTTCTCCCGGCTCAGGSCCAAGAAC 180
DB |||||
QY 181 GATGAGACAGCTGAGTGTGGGCGCAACACAGGATATCTGTGTAAAGCGTTCTCCCGG 240
DB |||||
QY 181 GATGAGACAGCTGAGTGTGGGCGCAACACAGGATATCTGTGTAAAGCGTTCTCCCGG 240
DB |||||
QY 241 CTCGGGCGCAAGACAGATGTCTCCAGATGGGTCCAGCCCTCAGCAGTTTCTAGTGAA 300
DB |||||
QY 241 CTCGGGCGCAAGACAGATGTCTCCAGATGGGTCCAGCCCTCAGCAGTTTCTAGTGAA 300
DB |||||
QY 301 TCATCAGATGTTTCCAGGCTGCCCAAGGACCTGAAATGACCTGTACCTTATTTGAAC 360
DB |||||
QY 301 TCATCAGATGTTTCCAGGCTGCCCAAGGACCTGAAATGACCTGTACCTTATTTGAAC 360
DB |||||
QY 361 TAACCAATCAGTTCGCTTCTCGCTTCTGTTCGGCGCTTCCGCTCTCCGAGCTCAATAA 420
DB |||||
QY 361 TAACCAATCAGTTCGCTTCTCGCTTCTGTTCGGCGCTTCCGCTCTCCGAGCTCAATAA 420
DB |||||
QY 421 AGAGCCCAACCCCTCACTCGGCGCGCAGTCTTCCGATAGACTGCGTCCCGGGTAC 480
DB |||||
QY 421 AGAGCCCAACCCCTCACTCGGCGCGCAGTCTTCCGATAGACTGCGTCCCGGGTAC 480
DB |||||
QY 481 CCGTATTTCCCAATAAAGCCTTCTGCTGTTTCATCCGATCGTGTCTGCTGTTCTTGG 540
DB |||||
QY 481 CCGTATTTCCCAATAAAGCCTTCTGCTGTTTCATCCGATCGTGTCTGCTGTTCTTGG 540
DB |||||
QY 541 GGAGGCTCTCTCTGAGTGTGATGATACCCACGACGGGGTCTTTCATTTGGGGCTCGT 600
DB |||||
QY 541 GGAGGCTCTCTCTGAGTGTGATGATACCCACGACGGGGTCTTTCATTTGGGGCTCGT 600
DB |||||
QY 601 CCGGATTTGAGACCCCTGCGCCAGGACACACGACCCACCGGAGGTAAAGTGGCC 660
DB |||||
QY 601 CCGGATTTGAGACCCCTGCGCCAGGACACACGACCCACCGGAGGTAAAGTGGCC 660
DB |||||
QY 661 AGCACTTATCTGTGCTCGGATGCTGATGCTAGTGTCTATGTTGATGTTATGCGCCTGG 720
DB |||||
QY 661 AGCACTTATCTGTGCTCGGATGCTGATGCTAGTGTCTATGTTGATGTTATGCGCCTGG 720
DB |||||
QY 721 TCTGTACTAGTTAGCTAAGTCTGATCTGCGGACCGCTGCTGAACTGACGAGTT 780
DB |||||
QY 721 TCTGTACTAGTTAGCTAAGTCTGATCTGCGGACCGCTGCTGAACTGACGAGTT 780
DB |||||
QY 781 CTGAACACCCCGCGCAACCCCTGGGAGACGTCACGAGGACTTTGGGGCCGTTTGTGG 840
DB |||||
QY 781 CTGAACACCCCGCGCAACCCCTGGGAGACGTCACGAGGACTTTGGGGCCGTTTGTGG 840
DB |||||
QY 841 CCGACCTGAGGAAGGAGTGCATGTGGAATCCGACCCCGCTCAGGATATGTTGTTGTT 900
DB |||||
QY 841 CCGACCTGAGGAAGGAGTGCATGTGGAATCCGACCCCGCTCAGGATATGTTGTTGTT 900
DB |||||
QY 901 AGGAGACGAGAACTTAAACAGTTTCCGCGCTCGGTCTGAAATTTTGTCTTCGTTTGGAA 960
DB |||||
QY 901 AGGAGACGAGAACTTAAACAGTTTCCGCGCTCGGTCTGAAATTTTGTCTTCGTTTGGAA 960
DB |||||
QY 961 CCGAAGCGCGCTCTTGTCTGCTGACGCAAGCTTGGGCTGCAGGCTCGAGGACTGGGGA 1020
DB |||||
QY 961 CCGAAGCGCGCTCTTGTCTGCTGACGCAAGCTTGGGCTGCAGGCTCGAGGACTGGGGA 1020
DB |||||
QY 1021 TCAATTGGCAGAGTAATCGGTGCTGCTGCTCTTTAGGACATATGAAATGAGTGCACAGT 1080
DB |||||
QY 1021 CC-----CTGCACCAAGAACATGAGAACTGAGAACTCAGGATTCCT 1059
DB |||||
QY 1081 GGGATGACTTTCTGATCAGCAAGAGGACATGACAGCTGTACAGAGTCTGTGAAGTTCG 1140
DB |||||
QY 1060 AGGACCCCTGCTGTGTTACAGGCGGGTCTTCTGTTGACAGAACTCTCAATACC 1119
DB |||||

QY 1141 ATGCTCGCTCAGTCACAGCTTTGCTTCTCTCCCATCTCTAAATAATGGCCCAACTCTTCAAG 1200
DB |||||
QY 1120 ACAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGAGCACCACTGTGCC 1179
DB |||||
QY 1201 AGAGGATGAAGCTCTTATAAARACTGACCTGATCACCTTTATCTCATTTGTTTGTAGTTC 1260
DB |||||
QY 1180 TGGCCAAAATTCGGAGTCCCAACCTCAATCATCACCAACCTCTTGTCTCTCCATTTG 1239
DB |||||
QY 1261 TCGTGGCCCATCATTTGGCATAGTGGCAGCTCAGCTCTCTGAAATGGGAAACGAGAAATGCA 1320
DB |||||
QY 1240 TCCTGGCTATC-----GCTGATGTGCTGCGCGTTTATC 1276
DB |||||
QY 1321 CGTTGGCTCAGTTAATGTCAGATATATCTCAGTCCGAAAGGCAAGAAATGGCAGTG 1380
DB |||||
QY 1277 ATATTCTCTTCACTCTGCTATGCTCATCTTCTTGTGTTCTTCTGGACTACCAA 1336
DB |||||
QY 1381 AAGATGAATGAGATTTTCGAGAACCTGTGATGGAACGATGAGCAACATGGAAGCAGAA 1440
DB |||||
QY 1337 GGTATGTTGCCGTTTGTCTCTTCTTCCAGGAA---CATCACTACAGCACCGGACCA 1393
DB |||||
QY 1441 TCAGATATCTTTCAGATAATGAAGCCAATCTCTAGATGCTAAGAAATTTCCAAAATTTCA 1500
DB |||||
QY 1394 TGCAAGACCTGCACGATTCCTGCTCAAGGAACCTCTATGTTTCCCTCTT-----GTTGCT 1448
DB |||||
QY 1501 GCATAACAACTGATCAAGATTTAATGATGTTTCTTTCAGCTAAAATTTCTTACTTCC 1560
DB |||||
QY 1449 GTACAAAACCTTCCGACGGAACCTGCACCTTGTATTCCTCATCCCATCTCTGGGCTTCG 1508
DB |||||
QY 1561 CCATCCAGGAACATGAGATATCATAGGGGATATCTCCAGTCTATTAGTGTCTGAACA 1620
DB |||||
QY 1509 CAAGATTCCTATGGAGTGGGCTCAGTCCGTTTCTCTCTG----- 1549
DB |||||
QY 1621 CCACAGTACTTGTATTTGTCAGTTCAGTATTTGAAACACATGAATGGCAGAGTCCAAAGAAATG 1680
DB |||||
QY 1550 -CTCAGTTACTAGTGCATTTGTTTCAGTGTTC---GTAGGCTTTTCCCACTGTTTG 1605
DB |||||
QY 1681 CATTTAAACAAAGAGGAGATGCGTAAATTTAGAGGAGCGTATATACATGATCATGACG 1740
DB |||||
QY 1606 GCTTTCAGTTATATG-GATGATGTTGTTTGGGGCCCAAGTCTGTACAACTCTTGAGTC 1664
DB |||||
QY 1741 AATTAAGTCTCTAGATGAACAAACAGTATATTTTGGAAACAGGAATTAAGGGGAAATGA 1800
DB |||||
QY 1665 CTTTTCACCTCTATTACCAATTTTCTTGTCTTGGGTATACATTTAAACCCCTTAATA 1724
DB |||||
QY 1801 AACCTGTTGAATAATATCACTAAATGATCTGAGGCTGAAGGATTTGGAAACATTTCTCAGACAT 1860
DB |||||
QY 1725 AAC-----CAACGT 1734
DB |||||
QY 1861 TGAATAATATCCTTTTACTCCAGGTGCCAGAAAGTCTCGCTGACTGGGAAATGGACCA 1920
DB |||||
QY 1735 TGGGGCTACTCCCTTAACTTCATGGGATATGTAATGGATGTTGGGG----- 1781
DB |||||
QY 1921 ACGATCTGGGCTCCAACTGACCATCGGGCTGTGAACAGCAGAGGTGAATTCACAGGCA 1980
DB |||||
QY 1782 ----- 1781
DB |||||
QY 1981 CCTACATCACAGCCGTAAACAGCCACATCAAAATGAGATCAAAAGAGTCCACCTGCATGGGA 2040
DB |||||
QY 1782 -----TACTTTACCGCAAGAACATATTTGTAATAAATAACAGCAATGTTTTCG- 1829
DB |||||
QY 2041 CACAAAACCAATCAACAGAGGACCCAGGCCACCTTTTGGCTTCCCGTCAATTTGGAAGT 2100
DB |||||
QY 1830 -----AAAACTGCTGTAATAGACCTTATGATTTGGAAGATGATGTACAGAGACTTGTGGGT 1884
DB |||||
QY 2101 TTTTCAGAGTCCACCATGCTTTCACGGGCCAGTGTCTTATAGACAGGAATGGGAGAGG 2160
DB |||||
QY 1885 CTTTGGGCTTGTGCTCCCTTTTACAAATGTGGCTATCTGCTTTAATGCTTTTAT 1944
DB |||||
QY 2161 TCCTGAAGACCATGTGCTGCTGCGGTCAAGTGTATGATGACATTTGGGTGATGACTGGAAAG 2220
DB |||||
QY 1945 GCATGATACAACT----- 1959
DB |||||
QY 2221 CTACAGGCTCGGCATCAACATCTTCACTCGCCTGCGCACACAGAGAGGTGAGTGAGTG 2280
DB |||||

Db 1960 -----AAGCAGGCTTCACTTCTCGCCAACTTACAGGCCCTTCTGTGTA 2005
Qy 2281 ACCAAGTCTCTCGGACTCCAGGTGAAAAAGGAGATAGAGCCCTCTCGACAAAATGG 2340
Db 2006 AACAAATATCTGAACCTTTTACCCCGTGTCCCGCA----- 2039
Qy 2341 TATACCAGGCTTTCAGGTCTCAATAGGTACTCCAGGTCTTAAAGTGTATCGGGGGATCT 2400
Db 2040 ----- 2039
Qy 2401 CTGTTTACCTGGAGTTCAGGATTTCCAGGACCAATGGGAAGACCGGAAGCCAGGAC 2460
Db 2040 ----- 2039
Qy 2461 TTAATGGACAAAAGGCCAGAGGAGAAAAAGGAGTGGAAAGCATGCAAAAGACAATCTA 2520
Db 2040 -----ACGGTCAG 2047
Qy 2521 ATACAGTCCGACTGTGGGTGCGAGCGCCCTCACGAAGGCAGAGTGGAGATTTTTCACG 2580
Db 2048 GTCTCTGCAAGTGTCTGACGCAACCCCACTGGAATGGGGCTTGGCTATCGGCCATA 2107
Qy 2581 AAGSCAGTGGGGTACGCTGTGTGACGACCGCTCGGAACCTCGTGGAGGACTGCTGCTCT 2640
Db 2108 GCCGATCGCGGACCTTTGTGGCTCTCTGCG----- 2141
Qy 2641 GCAGAGCTTGGGATACAAAGGTGTTCAAAGTGTGCATAGCGAGCTTATTTTGGAAAAAG 2700
Db 2142 -----ATCCATACTGGGGAACCTCTAGCAGCTTGT----- 2171
Qy 2701 GTACGGGTCCAATATGGCTGAATGAGATTAATTTTTCGGGAAGAGTCAATCCATTGAG 2760
Db 2172 ----- 2171
Qy 2761 AGTCAGAAATAGACAGTGGGGTGTGAGAGCTGTTGCGACGAGAAAGATGCTGGGGGTC 2820
Db 2172 -----TTTGCTCGCAGCGGCTGAGCGGAA 2197
Qy 2821 ACTTTGCACTACATAAATGCAATCATATTTTCAATCACAATTTTAAACTGTATAAAGTG 2880
Db 2198 ACTTATCGC----- 2207
Qy 2881 ATTTTTCCTTGTCTTCACTAABAACTAGCTTAATTAATATTTAAGRAACTAAGAAATTT 2940
Db 2208 ----- 2207
Qy 2941 ATCCAGAAAAAGGAATATTTAAAAATCACTGGATATAACATATAAAATAGCTTCATATTT 3000
Db 2208 -----ACCGACAACCTCT 2219
Qy 3001 GCTTCAAAATACAGAAACATTTTCAACTTCTCTAGGTTTTAAAGTGGCTCGTGGCGAAATG 3060
Db 2220 GTTGCTCTCTCGGAAATACACCTCTTTCATGTGCTGCTAGGGTGTGCTGCCAACTGG 2279
Qy 3061 ATCCCTCAGATATAGTATGTTTCGCTTTTCATAGGAGGGGGAATGTAGTCTTATGC 3120
Db 2280 ATCCCTCAGATATAGTATGTTTCGCTTTTGCATAGGAGGGGGAATGTAGTCTTATGC 2339
Qy 3121 AATACTCTTGTAGTCTTGCACATGGTAAACGATAGTTAGCAACATGCTTTTACAAGGAGA 3180
Db 2340 AATACACTGTAGTCTTGCACATGGTAAACGATAGTTAGCAACATGCTTTACAAGGAGA 2399
Qy 3181 GAAAAGCACCGTGCATCCGATTTGGTGGAAAGTAAAGTGTGATCGATCGTCTTATTAGG 3240
Db 2400 GAAAAGCACCGTGCATCCGATTTGGTGGAAAGTAAAGTGTGATCGATCGTCTTATTAGG 2459
Qy 3241 AAGGCAACAGAGGGTCTGCATGGAATGGAGGAACCACTGAATTCGCGATTCAGAGAT 3300
Db 2460 AAGGCAACAGAGGGTCTGCATGGAATGGAGGAACCACTGAATTCGCGATTCAGAGAT 2519
Qy 3301 -ATTGTATTTAAGTGCCTAGCTCGATACAGCAAAACGCCA--TTTGACCATTCACCAATTT 3357
|||||

Db 2520 AATTGTATTTAAGTGCCTAGCTCGATACAGCAAAACGCCATTTTGTGACCATTCACCAAT 2579
Qy 3358 GGTGTGCACCT--CCAAAGCTTTCAGCTGCGCAGCACTCAGGGCGCAAGGGCTGTCTAAA 3415
Db 2580 GGTGTGCACCTTCCAAAGCTTTCAGCTGCGCAGCACTCAGGGCGCAAGGGCTGTCTAAA 2639
Qy 3416 GGAAGCCGAAACAGGTAGAAAAGCCAGTCCGCAGAAAACGGTGTCTGACCCCGATGAATGTCA 3475
Db 2640 GGAAGCCGAAACAGGTAGAAAAGCCAGTCCGCAGAAAACGGTGTCTGACCCCGATGAATGTCA 2699
Qy 3476 GCTACTGGGCTATCTGGAACAAGGGAACAAACGCAAGCGCAAAAGAGAACAGAGTACTTGCA 3535
Db 2700 GCTACTGGGCTATCTGGAACAAGGGAACAAACGCAAGCGCAAAAGAGAACAGAGTACTTGCA 2759
Qy 3536 GTGGGCTTACATGCGGATAGCTAGACTGGGCGGTTTTTATGACAGCAAGCGAAACCGGAAT 3595
Db 2760 GTGGGCTTACATGCGGATAGCTAGACTGGGCGGTTTTTATGACAGCAAGCGAAACCGGAAT 2819
Qy 3596 TGGCAGCTGGGGCGCCCTCTGTGTAAGGTTGGGAAGCCCTGCAAAAGTAAACTGGATGGCTT 3655
Db 2820 TGGCAGCTGGGGCGCCCTCTGTGTAAGGTTGGGAAGCCCTGCAAAAGTAAACTGGATGGCTT 2879
Qy 3656 TCTTGGCGCAAGATCTGATGCGCAGGGGATCAAGATCTGATCAAGAGACAGAGATGAG 3715
Db 2880 TCTTGGCGCAAGATCTGATGCGCAGGGGATCAAGATCTGATCAAGAGACAGAGATGAG 2939
Qy 3716 GATCGTTTCCATGATGGAACAAGATGGAATGGAACGAGGTTCTCCGGCGCTTGGGGTG 3775
Db 2940 GATCGTTTCCATGATGGAACAAGATGGAATGGAACGAGGTTCTCCGGCGCTTGGGGTG 2999
Qy 3776 AGAGGCTATTTCGGCTATGACTGGGCACAACAGACAATCGGCTGTCTGTATGCCGCCGTGT 3835
Db 3000 AGAGGCTATTTCGGCTATGACTGGGCACAACAGACAATCGGCTGTCTGTATGCCGCCGTGT 3059
Qy 3836 TCCGGCTGTCAAGCGAGGGCGCCCGGTTCTTTTGTGAAGACCGACTGTGTCCGGTGCCC 3895
Db 3060 TCCGGCTGTCAAGCGAGGGCGCCCGGTTCTTTTGTGAAGACCGACTGTGTCCGGTGCCC 3119
Qy 3896 TGAATGAACCTGCAAGGACGAGCGGCTATCGTGGCTGGCCACGACCGGGCTTCCCTT 3955
Db 3120 TGAATGAACCTGCAAGGACGAGCGGCTATCGTGGCTGGCCACGACCGGGCTTCCCTT 3179
Qy 3956 GCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGAGTGTGCTGTATTTGGCGGAAG 4015
Db 3180 GCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGAGTGTGCTGTATTTGGCGGAAG 3239
Qy 4016 TGGCGGGGAGGATCTCTCTCATCTCACTTGTCTCGCGAGAAAGTATCCATCATGG 4075
Db 3240 TGGCGGGGAGGATCTCTCTCATCTCACTTGTCTCGCGAGAAAGTATCCATCATGG 3299
Qy 4076 CTGATGCAATGCGGGGCTGCATACGCTTGTATCGGCTACCTCGCCATTCGACCCCAAG 4135
Db 3300 CTGATGCAATGCGGGGCTGCATACGCTTGTATCGGCTACCTCGCCATTCGACCCCAAG 3359
Qy 4136 CGAAACATCCGATCGAGCAGCAGTACTCGGATGGAAGCGGCTCTTGTGTGATCAGGATG 4195
Db 3360 CGAAACATCCGATCGAGCAGCAGTACTCGGATGGAAGCGGCTCTTGTGTGATCAGGATG 3419
Qy 4196 ATCTGGAAGAGATCAGGGGCTCGCGCAGCGCAACTGTTCGCCAGGCTCAAGGGCG 4255
Db 3420 ATCTGGAAGAGATCAGGGGCTCGCGCAGCGCAACTGTTCGCCAGGCTCAAGGGCG 3479
Qy 4256 GCATGCCCGAGCGGAGGATCTGTGTGATGCCATGCGGATGCTGTGCTGCGGATATCA 4315
Db 3480 GCATGCCCGAGCGGAGGATCTGTGTGATGCCATGCGGATGCTGTGCTGCGGATATCA 3539
Qy 4316 TGGTGGAAATGGCGGCTTTTCTGGAATTCATCGACTGTGGCGGCTGGGTGTGGCGGACC 4375
Db 3540 TGGTGGAAATGGCGGCTTTTCTGGAATTCATCGACTGTGGCGGCTGGGTGTGGCGGACC 3599
Qy 4376 GCTATCAGGACATAGCGTGTGGCTACCGCTGATATGCTGAAGAGCTTTGGCGCGCAATGGG 4435
Db 3600 GCTATCAGGACATAGCGTGTGGCTACCGCTGATATGCTGAAGAGCTTTGGCGCGCAATGGG 3659

QY 4436 CTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTCGAGCGCATCGCCTTCT 4495
DB |||||
DB 3660 CTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTCGAGCGCATCGCCTTCT 3719
QY 4496 ATCGCTTCTTGACAGTCTCTCTGAGCGGACTCTCGGGTTCGATA 4542
DB |||||
DB 3720 ATCGCTTCTTGACAGTCTCTCTGAGCGGACTCTCGGGTTCGATA 3766

RESULT 9

ID ADM68972 standard; DNA; 5130 BP.
AC ADM68972;
XX
DT 03-JUN-2004 (first entry)
XX
DE LSRNL vector DNA.
XX
KW Pharmaceutical; industry; diagnosis; LSRNL; ds.
XX
OS Synthetic.
XX
PN US2003224415-A1.
XX
PD 04-DEC-2003.
XX
PF 26-MAR-2003; 2003US-00397079.
XX
PR 29-JUN-2001; 2001US-00897511.
PR 28-MAR-2002; 2002US-0368357P.
XX
PA (GALA-) GALA DESIGN INC.
XX
PI Brenel RD, Bleck GT, York D, Bakle K;
XX
DR WPI; 2004-033957/03.
XX

Host cell useful for producing proteins for pharmaceuticals, industrial purposes, comprising genome having at least one integrated vector having at least one exogenous gene and lacking selectable marker.

Example 1; SEQ ID NO 9; 78pp; English.

The invention relates to a host cell comprising a genome comprising at least one integrated integrating vector, where the integrating vector comprises at least one exogenous gene operably linked to a promoter, and where the integrating vector lacks a gene encoding a selectable marker. The invention also relates to a method of transfecting host cells, involving providing a number of host cells comprising a genome and a number of integrating vectors, where the integrating vectors comprise at least one exogenous gene, and where the integrating vectors lack a gene encoding a selectable marker, contacting the host cell with a number of integrating vectors to generate transfected host cells comprising at least one integrated copy of the integrating vector and selecting the transfected host cells. The host cell is useful for producing a protein of interest which involves providing a host cell, where the exogenous gene encodes a protein of interest and culturing the host cell under conditions such that the protein of interest is produced. The integrated exogenous gene is stable in the absence of selection. The integrating vector further comprises a secretion signal sequence operably linked to the exogenous gene. The host cell is useful for the production of proteins for pharmaceuticals and industrial, diagnostic and other purposes, and in the production of multiple variants of proteins, followed by analysis of the activity of the protein variants. This sequence represents LSRNL vector DNA, used in the scope of the invention.

Sequence 5130 BP; 1187 A; 1335 C; 1332 G; 1276 T; 0 U; 0 Other;

Query Match 30.7%; Score 1588.6; DB 12; Length 5130;

Best Local Similarity 66.1%; Pred. NO. 0;

Matches 3007; Conservative 0; Mismatches 754; Indels 786; Gaps 18;

QY 1 TTTGAAAGAGACCCCAACCGTAGGTGGCAAGCTAGCTTAAGTAAACGCCACTTTTGCAGGCAT 60
DB |||||
DB 1 TTTGAAAGAGACCCCAACCGTAGGTGGCAAGCTAGCTTAAGTAAACGCCACTTTTGCAGGCAT 60
QY 61 GGAATAATACATATCTAGATAGAAAAGTTTACATCAAGTTCAGGAAACAAAGAAACAGC 120
DB |||||
DB 61 GGAATAATACATATCTAGATAGAAAAGTTTACATCAAGTTCAGGAAACAAAGAAACAGC 120
QY 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTCTGCCCCCGGCTCAGGGCCCAAGAAC 180
DB |||||
DB 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTCTGCCCCCGGCTCAGGGCCCAAGAAC 180
QY 181 GATGAGACAGCTGATGATGGGCCAAACAGGATATCTGTGTAAGCGGTTCTGCCCCCGG 240
DB |||||
DB 181 GATGAGACAGCTGATGATGGGCCAAACAGGATATCTGTGTAAGCGGTTCTGCCCCCGG 240
QY 241 CTGGGGCCCAAGAACAGATGATGGTCCAGATGCGGTCCAGCCCTCAGCAGTTCTTAGTGA 300
DB |||||
DB 241 CTGGGGCCCAAGAACAGATGATGGTCCAGATGCGGTCCAGCCCTCAGCAGTTCTTAGTGA 300
QY 301 TCATCAGATGTTTCCAGGGTCCCAAGGACCTGAAAATGACCCCTGTACCTTTATTGAAC 360
DB |||||
DB 301 TCATCAGATGTTTCCAGGGTCCCAAGGACCTGAAAATGACCCCTGTACCTTTATTGAAC 360
QY 361 TAACCAATCAGTTCTGCTTCTGTGTCGCGCTTCCGCTCTCCGAGTCAATAAA 420
DB |||||
DB 361 TAACCAATCAGTTCTGCTTCTGTGTCGCGCTTCCGCTCTCCGAGTCAATAAA 420
QY 421 AGAGCCCAACACCCCTCACTCGGGCGCCAGTCTTCGATAGACTGCGTCCCGGGTAC 480
DB |||||
DB 421 AGAGCCCAACACCCCTCACTCGGGCGCCAGTCTTCGATAGACTGCGTCCCGGGTAC 480
QY 481 CCGTATTTCCCAATAAAGCCTCTTGTGTCGATCCGAATCGTGTCTCGCTGTTCTTGTG 540
DB |||||
DB 481 CCGTATTTCCCAATAAAGCCTCTTGTGTCGATCCGAATCGTGTCTCGCTGTTCTTGTG 540
QY 541 GGAGGGTCTCTCTGAGTGAATGATACCAACGAGGGGCTTTTCAATTTGGGGGCTCGT 600
DB |||||
DB 541 GGAGGGTCTCTCTGAGTGAATGATACCAACGAGGGGCTTTTCAATTTGGGGGCTCGT 600
QY 601 CCGGATTTGGAGACCCCTGCGGAGGACCAACCCAGCCACCGGGGAGTAAGCTGGCC 660
DB |||||
DB 601 CCGGATTTGGAGACCCCTGCGGAGGACCAACCCAGCCACCGGGGAGTAAGCTGGCC 660
QY 661 AGCAACTTATCTGTGTCGTCGATTTCTAGTCTCTATGTTTGTATGTCGCTCGC 720
DB |||||
DB 661 AGCAACTTATCTGTGTCGTCGATTTCTAGTCTCTATGTTTGTATGTCGCTCGC 720
QY 721 TCTGTACTAGTTAGTAACTAGCTCTGTATCTGCGGACCCGCTGGTGAACCTGACGAGTT 780
DB |||||
DB 721 TCTGTACTAGTTAGTAACTAGCTCTGTATCTGCGGACCCGCTGGTGAACCTGACGAGTT 780
QY 781 CTGAAACACCGCGCGCAACCCCTGGGAGACGTCACCGGACTTTGGGGGCGCTTTTGTGG 840
DB |||||
DB 781 CTGAAACACCGCGCGCAACCCCTGGGAGACGTCACCGGACTTTGGGGGCGCTTTTGTGG 840
QY 841 CCGGACCTGAGGAGGAGGAGTCTGATGCGAATCCGACCCCGCTCAGGATATGTTTGTGT 900
DB |||||
DB 841 CCGGACCTGAGGAGGAGGAGTCTGATGCGAATCCGACCCCGCTCAGGATATGTTTGTGT 900
QY 901 AGGAGACGAGAACCTTAAACAGTTTCCCGCTCTCGAATTTTGTCTTTCGTTTGGAA 960
DB |||||
DB 901 AGGAGACGAGAACCTTAAACAGTTTCCCGCTCTCGAATTTTGTCTTTCGTTTGGAA 960
QY 961 CCGAAGCGCGGCTCTTGTCTGTGACCAAGCTTGGGCTGACGCTGACACTAGAGGA 1020
DB |||||
DB 961 CCGAAGCGCGGCTCTTGTCTGTGACCAAGCTTGGGCTGACGCTGACACTAGAGGA 1020
QY 1021 TCAATTCGGACAGTAAATCGGTGCTGCGCTCTTTAGGACATATGAGTATGCACAGT 1080
DB |||||
DB 1021 CC-----CTGACCAGAACATGGAGAACACACATCAGGATTCCT 1059

Db 2460 AAGCAACAGACAGCTCTGACATGATTTGGAGCAACCACTGATTCGCCATTTCAGAGAT 2519
Qy 3301 -ATTGTAATTAAGTCCCTAGCTCGATACAGCAACGCCA--TTTGACCAATTCACACAT 3357
Db 2520 AATTGTAATTAAGTCCCTAGCTCGATACAGCAACGCCAATTTTGGACCAATTCACACAT 2579
Qy 3358 GGTGTGCACT--CCAAGCTTCACGCTCCGCAAGCACTCAGGGGCGCAAGGCTGCTAAA 3415
Db 2580 GGTGTGCACTTCCTAAAGCTTCACGCTCCGCAAGCACTCAGGGGCGCAAGGCTGCTAAA 2639
Qy 3416 GGAAGCGGAAACACGTAGAAAGCCAGTCCGCAAGAAAGCGTGTGACCCCGGATGAATGTCA 3475
Db 2640 GGAAGCGGAAACACGTAGAAAGCCAGTCCGCAAGAAAGCGTGTGACCCCGGATGAATGTCA 2699
Qy 3476 GCTACTGGGCTATCTGGCAAGAGGAAACGCAAGCGCAAGAGAAAGCAGGTAGCTTCA 3535
Db 2700 GCTACTGGGCTATCTGGCAAGAGGAAACGCAAGCGCAAGAGAAAGCAGGTAGCTTCA 2759
Qy 3536 GTGGGCTTACATGGCGATAGCTAGACTGGGCGGTTTATGACACAGCAAGCAAGCGGAT 3595
Db 2760 GTGGGCTTACATGGCGATAGCTAGACTGGGCGGTTTATGACACAGCAAGCAAGCGGAT 2819
Qy 3596 TGCCAGCTGGGCGGCGCTCTGTAGAGTTGGAGCCCTGCAAAAGTAACTGGATGGCTT 3655
Db 2820 TGCCAGCTGGGCGGCGCTCTGTAGAGTTGGAGCCCTGCAAAAGTAACTGGATGGCTT 2879
Qy 3656 TCTTGCCGCCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGACAGCAGATGAG 3715
Db 2880 TCTTGCCGCCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGACAGCAGATGAG 2939
Qy 3716 GATCGTTTCGATGATGAACAGATGATTTGCAGCAGGTTCTCCGCGCCTTGGGTGG 3775
Db 2940 GATCGTTTCGATGATGAACAGATGATTTGCAGCAGGTTCTCCGCGCCTTGGGTGG 2999
Qy 3776 AGAGCTATTCGGCTATGACTGGGCACAAAGCAATCGCTGCTCTCATGCCGCGGT 3835
Db 3000 AGAGCTATTCGGCTATGACTGGGCACAAAGCAATCGCTGCTCTCATGCCGCGGT 3059
Qy 3836 TCCGCTGTACGCGCAGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCGGGTGCC 3895
Db 3060 TCCGCTGTACGCGCAGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCGGGTGCC 3119
Qy 3896 TGAATGAATCTGACGACGAGGCGCGGCTATCTGGCTGGCCACGACGCGGCTTCCTT 3955
Db 3120 TGAATGAATCTGACGACGAGGCGCGGCTATCTGGCTGGCCACGACGCGGCTTCCTT 3179
Qy 3956 GCGCAGCTGCTCGACGCTTCTACTGAAGCGGGAAGGACTGGCTGCTATTTGGCGAAG 4015
Db 3180 GCGCAGCTGCTCGACGCTTCTACTGAAGCGGGAAGGACTGGCTGCTATTTGGCGAAG 3239
Qy 4016 TGCCGGGCGCAGGATCTCCTGTCTATCTCACCTTGCTCTCCGCGAGAAAGTATCCATCATGG 4075
Db 3240 TGCCGGGCGCAGGATCTCCTGTCTATCTCACCTTGCTCTCCGCGAGAAAGTATCCATCATGG 3299
Qy 4076 CTGATGCAATGCGGCGGTGATACGCTTGATTCGGGCTACTTGGCCATTCGACCAACCAAG 4135
Db 3300 CTGATGCAATGCGGCGGTGATACGCTTGATTCGGGCTACTTGGCCATTCGACCAACCAAG 3359
Qy 4136 CGAATCATGCTGACGACGACGACTACTCGGATGGAAGCGGCTTCTGCTCATGAGATG 4195
Db 3360 CGAATCATGCTGACGACGACGACTACTCGGATGGAAGCGGCTTCTGCTCATGAGATG 3419
Qy 4196 ATCTGGACGAAGAGCATCAGGGGCTCGCGCGCAGCGCAACTGTTGCCAGGCTCAAGGGCG 4255
Db 3420 ATCTGGACGAAGAGCATCAGGGGCTCGCGCGCAGCGCAACTGTTGCCAGGCTCAAGGGCG 3479
Qy 4256 GCATGCCCGCAGGCGAGGATCTGCTGTGACCCCATGCGCATGCTGCTTGGCGAATATCA 4315
Db 3480 GCATGCCCGCAGGCGAGGATCTGCTGTGACCCCATGCGCATGCTGCTTGGCGAATATCA 3539
Qy 4316 TGGTGGAAATGCGCGCTTTCTGATTCATGCACTGTGGCGGCTGGGTGTGGCGACC 4375
Db 3540 TGGTGGAAATGCGCGCTTTCTGATTCATGCACTGTGGCGGCTGGGTGTGGCGACC 3599

Qy 4376 GCTATCAGGACATAGCGTTGGCTACCGTGTATATTGCTGAAGAGCTTGGCGCGCAATGGG 4435
Db 3600 GCTATCAGGACATAGCGTTGGCTACCGTGTATATTGCTGAAGAGCTTGGCGCGCAATGGG 3659
Qy 4436 CTGACCGCTTCTCGTCTTTACGGTATCGCGCTCCCGATTCGACGCGCATCGCTTCT 4495
Db 3660 CTGACCGCTTCTCGTCTTTACGGTATCGCGCTCCCGATTCGACGCGCATCGCTTCT 3719
Qy 4496 ATCCGCTTCTTGACGAGTTCTTCTGAGCGGACTCTCGGGTTTCGATA 4542
Db 3720 ATCCGCTTCTTGACGAGTTCTTCTGAGCGGACTCTCGGGTTTCGATA 3766

RESULT 10

ADW28107
ID ADW28107 standard; DNA; 5130 BP.
XX
AC ADW28107;
XX DT 07-APR-2005 (first entry)
XX DE Nucleotide sequence of LSNRL vector.
XX KW recombinant protein; integrating vector; LSNRL vector; ss.
XX OS Synthetic.
XX FH Key
XX LTR Location/Qualifiers
FT 1..589
FT /*tag= a
FT /note= "MoMuSV 5'LTR"
FT 659..897
FT /*tag= b
FT /note= "retroviral packaging region"
FT 1034..1714
FT /*tag= c
FT /product= "Hepatitis B surface antigen"
FT 2279..2595
FT /*tag= d
FT /note= "RSV promoter"
FT 2951..3745
FT /*tag= e
FT /gene= "neomycin phosphotransferase gene"
FT 4537..5130
FT /*tag= f
FT /note= "MoMuSV 3'LTR"

WO2005007803-A2.

27-JAN-2005.

27-MAR-2003; 2003WO-US009325.

28-MAR-2002; 2002US-0368357P.

26-MAR-2003; 52US-00368357.

(GALA-) GALA DESIGN INC.

Brenel Robert D, York D, Bleck GT, Bakle KA;

WPI; 2005-102089/11.

New host cell comprising integrated integrating vector, used in the
absence of selection and for producing high levels of proteins.

Example 1; SEQ ID NO 9; 148pp; English.

The invention relates to the production of proteins in host cells. The
host cells contain multiple integrated copies of an integrating vector
comprising an exogenous gene operably linked to a promoter, and where the
integrating vector lacks a gene encoding a selectable marker. The
invention is used to produce proteins of interest at high levels. The

CC present sequence represents a LSURL vector. This vector is used to
CC produce proteins in the method of the invention.

XX Sequence 5130 BP; 1187 A; 1335 C; 1332 G; 1276 T; 0 U; 0 Other;

SQ Query Match 30.7%; Score 1588.6; DB 14; Length 5130;

Best Local Similarity 66.1%; Pred. No. 0;

Matches 3007; Conservative 0; Mismatches 754; Indels 786; Gaps 18;

```
Qy 1 TTTGAAAGACCCACCCAGCTAGGTGGCAAGCTAGCTTAAAGTAACGCCACTTTGCAAGGCAT 60
Dy |||||
Qy 1 TTTGAAAGACCCACCCAGCTAGGTGGCAAGCTAGCTTAAAGTAACGCCACTTTGCAAGGCAT 60
Dy |||||
Qy 61 GGAATAATACATACTAGTAAGATAGAAAGTTTCAGATCAAGCTCAGGACCAAGAAACAGC 120
Dy |||||
Qy 61 GGAATAATACATACTAGTAAGATAGAAAGTTTCAGATCAAGCTCAGGACCAAGAAACAGC 120
Dy |||||
Qy 121 TGAATACCAACACAGGATATCTGTGTAGCGGTTCTGCCCGGCTCAGGGCCCAAGAAC 180
Dy |||||
Qy 121 TGAATACCAACACAGGATATCTGTGTAGCGGTTCTGCCCGGCTCAGGGCCCAAGAAC 180
Dy |||||
Qy 181 GATGAGACAGTGTAGTGTAGGCGCAAAACAGGATATCTGTGTAGAGCATTCCTGCCCGG 240
Dy |||||
Qy 181 GATGAGACAGTGTAGTGTAGGCGCAAAACAGGATATCTGTGTAGAGCATTCCTGCCCGG 240
Dy |||||
Qy 241 CTCGGGCGCAAGACAGATGTCCTCCAGATGGGTCCAGCCCTCAGCAGTTTCTAGTGAA 300
Dy |||||
Qy 241 CTCGGGCGCAAGACAGATGTCCTCCAGATGGGTCCAGCCCTCAGCAGTTTCTAGTGAA 300
Dy |||||
Qy 301 TCATCAGATGTTTCCAGGGTCCCCNAGGACCTGAAATGACCCCTGTACCTTATTTGAAC 360
Dy |||||
Qy 301 TCATCAGATGTTTCCAGGGTCCCCNAGGACCTGAAATGACCCCTGTACCTTATTTGAA 360
Dy |||||
Qy 361 TAACCAATCAGTTCGCTTCTCGTCTCTGTTCCGCGCTTCCGCTCTCCGAGCTCAATAAA 420
Dy |||||
Qy 361 TAACCAATCAGTTCGCTTCTCGTCTCTGTTCCGCGCTTCCGCTCTCCGAGCTCAATAAA 420
Dy |||||
Qy 421 AGAGCCCAACACCTCTACTCGGCGCGCAGTCTTCGATAGACTGCGTCCCGGGTAC 480
Dy |||||
Qy 421 AGAGCCCAACACCTCTACTCGGCGCGCAGTCTTCGATAGACTGCGTCCCGGGTAC 480
Dy |||||
Qy 481 CCGTATTCCTCAATAAGCCTCTTGCTGTTTGATCCGAACTCGTCTCGCTGTTCCCTTG 540
Dy |||||
Qy 481 CCGTATTCCTCAATAAGCCTCTTGCTGTTTGATCCGAACTCGTCTCGCTGTTCCCTTG 540
Dy |||||
Qy 541 GGAGGGTCTCTCTGAGTGAATGACTACCCACGAGCGGGTCTTTTCATTTGGGGGCTCGT 600
Dy |||||
Qy 541 GGAGGGTCTCTCTGAGTGAATGACTACCCACGAGCGGGTCTTTTCATTTGGGGGCTCGT 600
Dy |||||
Qy 601 CCGGATTTGGAGACCCCTGCCAGGACACCCAGACCCACCCAGGGAGGTAAAGTGGCC 660
Dy |||||
Qy 601 CCGGATTTGGAGACCCCTGCCAGGACACCCAGACCCACCCAGGGAGGTAAAGTGGCC 660
Dy |||||
Qy 661 AGCAACTTATCTGTCTGTCCGATTGTCTAGTGTCTATGTTGATGTTATGCGCCTCGC 720
Dy |||||
Qy 661 AGCAACTTATCTGTCTGTCCGATTGTCTAGTGTCTATGTTGATGTTATGCGCCTCGC 720
Dy |||||
Qy 721 TCTGTACTAGTAACTAGTCTGTATCTGCGGACCCCGTGGTGAACCTGACGAGTT 780
Dy |||||
Qy 721 TCTGTACTAGTAACTAGTCTGTATCTGCGGACCCCGTGGTGAACCTGACGAGTT 780
Dy |||||
Qy 781 CTGAACACCCCGCCCAACCTTGGGAGACTCCAGGACCTTGGGGCCGTTTTTGTGG 840
Dy |||||
Qy 781 CTGAACACCCCGCCCAACCTTGGGAGACTCCAGGACCTTGGGGCCGTTTTTGTGG 840
Dy |||||
Qy 841 CCCGACCTGAGGAAGGAGTGTGATGTGAAATCCGACCCCGCTCAGGATATGTTGTTCTGGT 900
Dy |||||
Qy 841 CCCGACCTGAGGAAGGAGTGTGATGTGAAATCCGACCCCGCTCAGGATATGTTGTTCTGGT 900
Dy |||||
Qy 901 AGGAGACGAGAACCTTAAACAGTTTCCCGCCTCCGCTCTGAATTTTTCGTTTGGAA 960
Dy |||||
Qy 901 AGGAGACGAGAACCTTAAACAGTTTCCCGCCTCCGCTCTGAATTTTTCGTTTGGAA 960
Dy |||||
```

```
Qy 961 CCGAAGCGCGCGCTTGTCTGTGACCAAGCTTGGGCTGCGAGTCTGACTCTAGAGA 1020
Dy |||||
Qy 961 CCGAAGCGCGCGCTTGTCTGTGACCAAGCTTGGGCTGCGAGTCTGAGGACTGGGA 1020
Dy |||||
Qy 1021 TCAATTCGGCACGAGTAAATCGGTGCTGCTCTTTAGGACATATGAAGTATGGACAGT 1080
Dy |||||
Qy 1021 CC-----CTGCACCGAACATGGAGAACACACATCAGGATTCCT 1059
Dy |||||
Qy 1081 GGGATGATTTTCTGTGATCAGACAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTCG 1140
Dy |||||
Qy 1060 AGGACCCCTCTCGTGTACAGCGGGGTTTTCTTGTGACAGAAATCTTCAACAATCC 1119
Dy |||||
Qy 1141 ATGCTCGCTCAGTGACAGCTTTGCTCTCCCAATCTTAAATATGAGGAGGAGGAGGAGG 1200
Dy |||||
Qy 1120 ACAGAGTCTAGACTCGGTGGACTTCTCAATTTTCTAGGGGAGGAGGAGGAGGAGG 1179
Dy |||||
Qy 1201 AGAGGATGAAGTCTTATTAATAATGCACTGATCAACCTTTTATCTCATTTGTGTGTAGTTC 1260
Dy |||||
Qy 1180 TGGCAAAATTCGCGAGTCCCAACCTCAATCACTACCAACCTCTTGTCTCTCAATTTG 1239
Dy |||||
Qy 1261 TCGTGCCCATCATTTGGCATAGTGGCAGCTCAGCTCTGAAATGGGAAACGAGAAATTGCA 1320
Dy |||||
Qy 1240 TCCTGGCTATC-----GCTGGATGTGTCTGCGGCGTTTTATC 1276
Dy |||||
Qy 1321 CGGTTGGCTCAGTTAAATGACAGATATATCTCAAGTCCGGAAGGCAAAAGGAAATGGCAGTG 1380
Dy |||||
Qy 1277 ATATTCCTCTTCACTCTGCTATGCTCATCTTCTTGTGTGTCTTCTGGACTACCA 1336
Dy |||||
Qy 1381 AAGATGAATAGATTTGAGAGAGCTGTGAGAACGATGAGCAACATGAGAAACATGAAAGCAGAA 1440
Dy |||||
Qy 1337 GGTATGTTGCGCGTTTGTCTCTACTTCCAGGAA--CATCAACTACAGCAGCGGACCA 1393
Dy |||||
Qy 1441 TCCAGTATCTTTTCAGATAATGAAGCCAAATCTCTAGATGCTAAGAAATTTCCAAATTTCA 1500
Dy |||||
Qy 1394 TGCAAGACCTGCGAGATCTCTGCTCAAGGAACCTCTATGTTTCCCTCTT-----GTTGCT 1448
Dy |||||
Qy 1501 GCATAACAACTGATCAAGATTTAAATGATGTTCTTTTCCAGTAAATTCCTTACTTTCTCCT 1560
Dy |||||
Qy 1449 GTACAAACCTTCGGACGGAAACCTGCACTTGATTTCCATCCCATCATCTCTGGGCTTTCG 1508
Dy |||||
Qy 1561 CCATCCAGGAACATGAGATATCATAGGGATATCTCCAAGTCAATTAAGTCTGAACA 1620
Dy |||||
Qy 1509 CAAGATTCCTATGGAGTGGGCTCAGTCCGTTCTCTCTCG----- 1549
Dy |||||
Qy 1621 CCAGAGTACTTGATTTGCACTTCAAGTATGAAACATCTGAATGGCAGAGTCCAGAGAAATG 1680
Dy |||||
Qy 1550 -CTCAGTTTACTAGTGCATTTGTTTCAAGTGTTC---GTAGGGCTTTCCCGCACTGTTG 1605
Dy |||||
Qy 1681 CATTTAAACAAACAGAGGAGATCGTAAATTTAGAGGAGCGTATATACAATGCAATGCAATGAG 1740
Dy |||||
Qy 1606 GCTTTCAAGTTATG-GATGATGTTGTTTGGGGGCCAAGTCTGTACACATCTTGAGTC 1664
Dy |||||
Qy 1741 AAATTAAGTCTCTAGATGAAACAAAGTATATTTGGAAACAGGAAATAAAAGGGGAAATGA 1800
Dy |||||
Qy 1665 CTTTTTACCTCTATTACCAATTTTCTTTGTTCTTGGGTATACATTTAAACCCCTAATAA 1724
Dy |||||
Qy 1801 AACTGTTGAATATATACATTAAGATCTGAGGCTGAAGGATGGGAACATTTCTCAGACAT 1860
Dy |||||
Qy 1725 AAC-----CAAAAGT 1734
Dy |||||
Qy 1861 TGAATAATATACATTTTACTCCAGGTGCCAGAAAGTCTCGCTGACTGGGAAATGGACCA 1920
Dy |||||
Qy 1735 TGGGGCTACTCCCTTAACTTCATGGGATATGTAATTTGATGTTGGG----- 1781
Dy |||||
Qy 1921 ACGATCTGGGCTCCAAATGACCATCGGGGCTGTGAACAGCAGAGGAGTGAATTCACAGCA 1980
Dy |||||
Qy 1782 ----- 1781
Dy |||||
Qy 1981 CCTATACACAGCGGTAAAGCCCATCAATAGATGATCAAGAGTCAACCACTGATGGGA 2040
Dy |||||
Qy 1782 -----TACTTTTACCGCAAGAACATATTTGTAATAAATAAATCAAGCAATTTTTCG- 1829
Dy |||||
Qy 2041 CACAAACACCATCAACAGAGGAGGCCAGCCACCTTTTGGCTTCCACCGTCAATTTGGAAGT 2100
Dy |||||
```


QY 4256 GCATGCCGACGGCAGGATCTGTCGTGACCCATGGCGATGCTTGTGCGGAATATCA 4315
DB |||||||
QY 3480 GCATGCCGACGGCAGGATCTGTCGTGACCCATGGCGATGCTTGTGCGGAATATCA 3539
DB |||||||
QY 4316 TGGTGGAAATGGCGCTTTTCTGGAATTCATCGACTGTGGCGGCTGGGTGTGCGGAC 4375
DB |||||||
QY 3540 TGGTGGAAATGGCGCTTTTCTGGAATTCATCGACTGTGGCGGCTGGGTGTGCGGAC 3599
DB |||||||
QY 4376 GCTATCAGGACATAGCTTGGCTACCCGTGATATTCCTGAAGAGCTTGGCGGCAATGGG 4435
DB |||||||
QY 3600 GCTATCAGGACATAGCTTGGCTACCCGTGATATTCCTGAAGAGCTTGGCGGCAATGGG 3659
DB |||||||
QY 4436 CTGACCGCTTCTCTGTGCTTTTACGGTATCGCGCTCCCGATTCCGAGCGCATCGCTTCT 4495
DB |||||||
QY 3660 CTGACCGCTTCTCTGTGCTTTTACGGTATCGCGCTCCCGATTCCGAGCGCATCGCTTCT 3719
DB |||||||
QY 4496 ATCGCTTCTTACGAGTTCCTTCTGAGCGGACTCTGGGGTTCGATA 4542
DB |||||||
QY 3720 ATCGCTTCTTACGAGTTCCTTCTGAGCGGACTCTGGGGTTCGATA 3766
DB |||||||

RESULT 11

AEB77874

ID AEB77874 standard; DNA; 5130 BP.

XX AC AEB77874;

XX AC AEB77874;

XX DT 20-OCT-2005 (first entry)

XX DE Retroviral vector expressing hepatitis B surface antigen.

XX DE Vector; gene therapy; cell transduction; transfection;

XX KW protein production; ds; surface antigen.

XX KW Moloney murine sarcoma virus.

XX OS Respiratory syncytial virus.

XX OS Moloney murine leukemia virus.

XX OS Hepatitis B virus.

XX OS Synthetic.

XX OS Unidentified.

XX Key Location/Qualifiers

FH LTR 1. 589

FT /tag= a

FT /note= "MoMuSV 5' LTR"

FT 659. .897

FT /tag= b

FT /note= "Retroviral packaging region"

FT 1034. .1714

FT /tag= c

FT /product= "hepatitis B surface antigen"

FT 2279. .2595

FT /tag= d

FT /note= "RSV promoter"

FT 2951. .3745

FT /tag= e

FT /note= "Neomycin phosphotransferase gene"

FT 4537. .5130

FT /tag= f

FT /note= "MoMuLV 3' LTR"

FT 4537. .5130

FT /tag= g

FT /note= "MoMuLV 3' LTR"

FT 4537. .5130

FT /tag= h

FT /note= "MoMuLV 3' LTR"

FT 4537. .5130

FT /tag= i

FT /note= "MoMuLV 3' LTR"

FT 4537. .5130

FT /tag= j

FT /note= "MoMuLV 3' LTR"

FT 4537. .5130

FT /tag= k

FT /note= "MoMuLV 3' LTR"

FT 4537. .5130

FT /tag= l

FT /note= "MoMuLV 3' LTR"

FT 4537. .5130

FT /tag= m

FT /note= "MoMuLV 3' LTR"

XX

DR

XX

XX

PT

PT

PT

XX

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Query Match

Best Local Similarity

Matches 3007;

Conservative

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

WPI; 2005-555595/56.

Transducing host cells useful for producing essential proteins comprises serially contacting host cells with integrating retroviral vectors to provide host cells comprising multiple integrated retroviral vectors.

Example 1; SEQ ID NO 9; 168pp; English.

The invention relates to transducing host cells comprising providing at least one host cell comprising a genome and a population of retroviral vectors encoding a gene of interest, contacting the host cell with the population of integrating vectors (under conditions such that the host cells are transduced to produce transduced host cells) and repeating both steps several times (e.g. at least 4-20) to provide host cells comprising multiple integrated retroviral vectors (about 10-100 integrated retroviral vectors). Also included is the host cell produced by the above method. The genes of interest are arranged in a polyclonistic sequence, and comprise immunoglobulin heavy and light chains. The integrating vector further comprises a secretion signal sequence operably linked to the exogenous gene. The retroviral vector further encodes an amplifiable marker selected from dihydrofolate reductase (DHFR) and glutamine synthetase. The method is useful for producing host cells that express recombinant proteins, which may be essential to the therapeutic treatment of many diseases and conditions. The present sequence is retroviral vector of the invention expressing hepatitis B surface antigen.

Sequence 5130 BP; 1187 A; 1335 C; 1332 G; 1276 T; 0 U; 0 Other;

Query Match 30.7%; Score 1588.6; DB 14; Length 5130;

Best Local Similarity 66.1%; Pred. No. 0;

Matches 3007; Conservative 0; Mismatches 754; Indels 786; Gaps 18;

1 TTTGAAAGACCCACCGTAGGTGGCAAGCTAGCTTAAGTAAACCCACTTTGCAAGGCAT 60

1 TTTGAAAGACCCACCGTAGGTGGCAAGCTAGCTTAAGTAAACCCACTTTGCAAGGCAT 60

61 GGAATAATACATACTGAGATAGAAAGTTTCAGATCAAGTTCAGGAAACAAAGAACAGC 120

61 GGAATAATACATACTGAGATAGAAAGTTTCAGATCAAGTTCAGGAAACAAAGAACAGC 120

121 TGAATACAAACAGGATATCTGTGTAAGCGGTTCTCTGCCCGCTCAGGCGCCAGAACAA 180

121 TGAATACAAACAGGATATCTGTGTAAGCGGTTCTCTGCCCGCTCAGGCGCCAGAACAA 180

181 GATGAGACAGCTGAGTATGCGGCAACACAGATATCTGTGTAAGCGGTTCTCTGCCCGG 240

181 GATGAGACAGCTGAGTATGCGGCAACACAGATATCTGTGTAAGCGGTTCTCTGCCCGG 240

241 CTGCGGCGCAAGAACAGATGCTCCCGAGATGCGGTCAGGCGCTCAGCAGTTTCTAGTGAA 300

241 CTGCGGCGCAAGAACAGATGCTCCCGAGATGCGGTCAGGCGCTCAGCAGTTTCTAGTGAA 300

301 TCATCAGATGTTTCCAGGGTGCCTCCAGAGACCTGAAATGACCTGTACCTTATTGAAC 360

301 TCATCAGATGTTTCCAGGGTGCCTCCAGAGACCTGAAATGACCTGTACCTTATTGAAC 360

361 TAACCAATCAGTTCGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 420

361 TAACCAATCAGTTCGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 420

421 AGAGCCCAACACCCCTCCTCCTGCGGCGGCTTCTCCGATAGACTGCTGCTGCTGCTGCTGCT 480

421 AGAGCCCAACACCCCTCCTCCTGCGGCGGCTTCTCCGATAGACTGCTGCTGCTGCTGCTGCT 480

481 CCGTATTCCCAATAAAGCCCTCTTGTCTGTTGTCATCCGAATCGTGGTCTCGCTGTTCTTGT 540

481 CCGTATTCCCAATAAAGCCCTCTTGTCTGTTGTCATCCGAATCGTGGTCTCGCTGTTCTTGT 540

541 GGAGGGTCTCTCTGAGTGTGATGATACCCAGAGCGGGGTCTTTCAATTTGGGGGCTCGT 600

541 GGAGGGTCTCTCTGAGTGTGATGATACCCAGAGCGGGGTCTTTCAATTTGGGGGCTCGT 600

QY 601 CCGGATTTGGAGACCCCTGCCAGGACCCAGCCACCCAGGAGGTAAGCTGCCC 660
Db 601 CCGGATTTGGAGACCCCTGCCAGGACCCAGCCACCCAGGAGGTAAGCTGCCC 660
QY 661 AGCAACTTATCTGTCTGTCCGATGCTAGTGTCTATGTTGATGTTATGCGCCTCG 720
Db 661 AGCAACTTATCTGTCTGTCCGATGCTAGTGTCTATGTTGATGTTATGCGCCTCG 720
QY 721 TCTGTACTAGTTAGTAACTAGCTCTGTATCTGGCGGACCCGTTGGAACTGACGAGTT 780
Db 721 TCTGTACTAGTTAGTAACTAGCTCTGTATCTGGCGGACCCGTTGGAACTGACGAGTT 780
QY 781 CTGAACACCCGCGCCCAACCTTGGAGAGCTCCAGGAGACTTTGGGCGCGTTTGTGG 840
Db 781 CTGAACACCCGCGCCCAACCTTGGAGAGCTCCAGGAGACTTTGGGCGCGTTTGTGG 840
QY 841 CCGGACCTTGAGGAGGAGTGCATCTGGAATCCGACCCCGCTCAGGATATGTTCTGCT 900
Db 841 CCGGACCTTGAGGAGGAGTGCATCTGGAATCCGACCCCGCTCAGGATATGTTCTGCT 900
QY 901 AGGAGACGAGAACTTAAACAGTTCCTCCGCTCCGCTCGTCTGAATTTTTCGTTTGGAA 960
Db 901 AGGAGACGAGAACTTAAACAGTTCCTCCGCTCCGCTCGTCTGAATTTTTCGTTTGGAA 960
QY 961 CCGAAGCCGCGCTTGTCTGCTGCAGCAGCTTGGGCTGCAGCTCTAGAGGA 1020
Db 961 CCGAAGCCGCGCTTGTCTGCTGCAGCAGCTTGGGCTGCAGCTCTAGAGGA 1020
QY 1021 TCAATTCGGCAGAGTAAATCGGTGCTGCGCTTTTAGACATATGAAGTATGCACAGT 1080
Db 1021 CC-----CTGCACCGAACATGGAGAACACACATCAGGAATTCCT 1059
QY 1081 GGGATGACTTTCTGATCAGCAAGAGGACACTGCAGCTGTACAGAGTCTGTGAAGTTTCG 1140
Db 1060 AGGACCCCTGCTCGTGTACAGCGGGTCTTTCTGTTGACAGAACTCTCACAATACC 1119
QY 1141 ATGCTGCTCAGTGACAGCTTTGCTCTCCCATCTTAAATAAGGCGCAACTCTTCAAG 1200
Db 1120 ACAGAGTCTAGACTCGTGTGAGTCTCTCAATTTCTAGGGGAGCACCACTGTCTCC 1179
QY 1201 AGAGGATGAAGTCTTATAAACTGCATGATCACCCTTTATCTCATTTGTTGTAGTTTC 1260
Db 1180 TGGCCAAATTCGGAGTCCCAACTCCTCAATCACTCACCACCTCTTGTCTCCAATTTG 1239
QY 1261 TCGTGCCCATCATTTGGCATAGTGCGAGCTCAGCTCTGAAATGGGAAACGAAGATTGCA 1320
Db 1240 TCCTGGCTATC-----GCTGGATGTCTCGGCGTTTATC 1276
QY 1321 CGGTTGGCTCAGTTAATCAGATATATCTCAAAGTCCGGAAGGCAAGAAATGGCAGTG 1380
Db 1277 ATATTCCTCTTCATCCTGCTATGCTCATCTCTTCTGTTGTTCTTCTGAGCTACCAA 1336
QY 1381 AAGATGAATAGATTTCCGAGAGCTGTGGAACGATGAGCAACATGGAAGCAGAA 1440
Db 1337 GGTATGTTGCCGTTTGTCTCTACTTCCAGGAA---CATCACTTACAGCACCGGACCA 1393
QY 1441 TCCAGTATCTTTCAGATAATGAAGCAATCTCTAGATGCTAAGAAATTTCCAAATTTCA 1500
Db 1394 TGCAGACTGCAGATTCCTGCTCAAGAACTCTATGTTTCCCTCTT-----GTTGCT 1448
QY 1501 GCAATCAACTGATCAAAAGATTAATGATGTTCTTTTCCAGCTAAATCTTACTTTCCCT 1560
Db 1449 GTACAAACCTTGGACGGAACCTGCACTGTATTTCCCATCCATCATCTCGGGCTTTCG 1508
QY 1561 CCATCCAGGAACATGAGATATCATAGGGGATATCTCCAAGTCAATAGTAGTCTGAACA 1620
Db 1509 CAAGATTCCTATGGAGTGGGCTCAGTCCGTTCTCTCTCG----- 1549
QY 1621 CCACAGTACTTGTATTTGAGTTTCACTATTTGAAACACTGAATGGCAGAGTCCAAGAGATG 1680
Db 1550 -CTCAGTTTACTAGTGCCATTTGTTTCACTGTTTC---GTAGGGCTTTCCCACTGTTT 1605
QY 1681 CATTTAAACAAAGAGGAGATGCGTAAATTTAGAGGAGCGTATATACAAATGCATCAGCAG 1740

Db 1606 GCCTTCAGTTATATG-GATGATGTGGTATTGGGGGCCCAAGTCTGTACAACTCTTGAGTC 1664
QY 1741 AAATTAAGTCTCTAGATGAAAAACAAGTATATTTTGGAAACAGAAATAAAAGGGAATGA 1800
Db 1665 CTTTTTACCTCTATACCATTCTTTTCTTTTGGGTATATCATTTAAACCTTAAATA 1724
QY 1801 AACTGTTGAATAATATCATTAATGATCTGAGGCTGGAAGGATTTGGGAACATTTCTCAGACAT 1860
Db 1725 AAC-----CAAGCT 1734
QY 1861 TGAATAATATCATTTTACTTCAAGGTGCCAGAAAGTGTCTGCTGACTGGGAAATGGACA 1920
Db 1735 TGGGGCTACTCCCTTAACTTCATGGGATATGTAATTTGGATGTTGGG----- 1781
QY 1921 ACATGCTGGGCTCAACATGACCATCGGGGCTGTGAACAGCAGAGGTGAATTCACAGGCA 1980
Db 1782 ----- 1781
QY 1981 CCTACATCAGAGCGTAAACAGCCACATCAAAATGAGATCAAAAGATCAACCTGCATGGGA 2040
Db 1782 -----TACTTTACCGCAAGAACATATTGTACTTAAAAATCAAGCAATGTTTCG- 1829
QY 2041 CACAAAACACCATCAACAGAGGACCCAGCCACCTTTGGCTTCAACCTCAATTTGGAAGT 2100
Db 1830 -----AAACCTGCTGTAAATAGACCTATTGATTTGGAAGATGTGTGAGAGCTTTGTGGT 1884
QY 2101 TTTCCAGAGTCCACCACTGTCTTTCAGCGGCCAGTCTCATAGACAGGAATGGGAAGGAGG 2160
Db 1885 CTTTTGGGCTTGTGCTGCTTTTACAAATGTGGCTATCTGCTCTTAATGCTTTATAT 1944
QY 2161 TCCTGAAGACATGTGGCTGCTGCGGTCAAGTGTTAATGACATTTGGTGTGATGATGGAAG 2220
Db 1945 GCATGTATACATCT----- 1959
QY 2221 CTACAGGCTCGGCATCAACATCTTCACTGCTGCGCACACAGAGAGTGTAGTGTG 2280
Db 1960 -----AAGCAGGCTTTTCACTTTCTCGCCAACTTACAAGGCTTTCTGTGTA 2005
QY 2281 ACCAAGTCTCTCGGACTCCAGGTGAAAGGAGATAGAGCCCTCTCGACAAAAATGG 2340
Db 2006 AACAAATCTGAACTTTTACCCCGTTGCGCGCA----- 2039
QY 2341 TATACCAGGCTTTCCAGGTCTAATAGTACTTCCAGGTCTTAAAGGTGATCGGGGGATCT 2400
Db 2040 ----- 2039
QY 2401 CTGTTTACCTGGAGTTGAGGATTTCCAGGACCAATGGGGAAGACCGGGAAGCCAGGAC 2460
Db 2040 ----- 2039
QY 2461 TTAATGGACAAAAGGCCAGAGGAGGAGAAAGGGAGTGGAGCATGCAAGACAACTCA 2520
Db 2040 -----ACGTCAG 2047
QY 2521 ATACAGTCCGACTGGTGGTGGCAGCGCCCTCAGAGGAGGAGTGGAGATTTTTCACG 2580
Db 2048 GTCTCTGCCAAGTGTGTGTGACCAACCCCACTGGATGGGCTTGGCTATCGGCCATA 2107
QY 2581 AAGGCCAGTGGGGTACGGTGTGTGACGACCGCTGGGAACTGCGTGGAGGACTGGTCTGCT 2640
Db 2108 GCCGCATCGCGGACCTTTTGTGGCTCTCTGCG----- 2141
QY 2641 GCAGGAGCTTGGGATACAAAGGTGTTCAAGTGTGCATTAAGCGAGCTTATTTTGGAAAAG 2700
Db 2142 -----ATCCATCTCGGAACTCTTAGCAGCTTGT----- 2171
QY 2701 GTACGGTCCAAATATGCTGAATGAAGTATTTTCTTTCGGGAAAGAGTCACTCCATTTGAAG 2760
Db 2172 ----- 2171
QY 2761 AGTCAGAAATPAGACATGGGGTGTGAGAGCTGTTCGACGACGAGAAAGATCTCGGGGTC 2820

Db 2172 -----TTTGTCTCGCAGCGGTCTGGAGCGAA 2197
Qy 2821 ACTTTGACCTACATAAATGCATCATATTTTCATTACATTTTTTAAACTGTTATAAAGTG 2880
Db 2198 ACTTATCGGC----- 2207
Qy 2881 ATTTTTCCTTTGCTTCACTAAATCAGCTTAATTAATATTTAAGAAACTAAGAAATTTT 2940
Db 2208 ----- 2207
Qy 2941 ATCCACAGAAAGGAATATTTTAAATAATCACTGGATAAACAATATAAATAGCTTCATATTT 3000
Db 2208 -----ACCGACACTCT 2219
Qy 3001 GCTTCAATACAGAACCAATTTCAACTTCTCTAGGTTTTTAAGTGGCTCGTGCAGAAATG 3060
Db 2220 GTTGTCTCTCTCGGAAATACACTCTCTTTCATGGCTGCTAGGCTGTGCTGCCAATCGG 2279
Qy 3061 ATCCCTCAGGATATAGTAGTTTCGCTTTTGCATAGGAGGGGAAATGTAGTCTTTATGC 3120
Db 2280 ATCCCTCAGGATATAGTAGTTTCGCTTTTGCATAGGAGGGGAAATGTAGTCTTTATGC 2339
Qy 3121 AATACTCTTAGTCTTCAACATAGTTAAACATAGTTAGCAACATGCTTTACAGGAGA 3180
Db 2340 AATACACTTGTAGTCTTCAACATAGTTAAACATAGTTAGCAACATGCTTTACAGGAGA 2399
Qy 3181 GAAAAAGCACCGTGCATCGCATTTGGTGGAAAGTAGGTGATACGATCGTGCCTTTATTAGG 3240
Db 2400 GAAAAAGCACCGTGCATCGCATTTGGTGGAAAGTAGGTGATACGATCGTGCCTTTATTAGG 2459
Qy 3241 AAGCAACAGACGGTCTGACATGATTTGGACGCAACCACTGAATTTCCGCAATTCAGAGAT 3300
Db 2460 AAGCAACAGACGGTCTGACATGATTTGGACGCAACCACTGAATTTCCGCAATTCAGAGAT 2519
Qy 3301 -ATTGTATTTAAGTCCCTAGCTCGATACAGCAAAAGCGCA--TTTGACCAATTTACACCAATT 3357
Db 2520 AATTGTATTTAAGTCCCTAGCTCGATACAGCAAAAGCGCAATTTTGGACCAATTTACACCAATT 2579
Qy 3358 GGTGTGCACCT--CCAAAGCTTCAGCTCGCAGCAAGCACTCAGGGCGCAAGGGCTGCTAAA 3415
Db 2580 GGTGTGCACCTTTCCAAAGCTTCAGCTCGCAGCAAGCACTCAGGGCGCAAGGGCTGCTAAA 2639
Qy 3416 GGAAGCGGAACAAGTAGAAGCCAGTCCGCGAGAAACGGTGTGACCCCGGATGAATGTCA 3475
Db 2640 GGAAGCGGAACAAGTAGAAGCCAGTCCGCGAGAAACGGTGTGACCCCGGATGAATGTCA 2699
Qy 3476 GCTACTGGGCTATCTGGACAAGGGGAAACCGCAAGCGCAAAAGAGAGAGAGTAGTTGCA 3535
Db 2700 GCTACTGGGCTATCTGGACAAGGGGAAACCGCAAGCGCAAAAGAGAGAGAGTAGTTGCA 2759
Qy 3536 GTGGGCTTACATGGGATAGCTAGACTGGGCGGTTTTATGACAGCAAGCGCAACCGGAAT 3595
Db 2760 GTGGGCTTACATGGGATAGCTAGACTGGGCGGTTTTATGACAGCAAGCGCAACCGGAAT 2819
Qy 3596 TGCCAGCTTGGGCGCCCTCTGTTAAGTTGGGAAGCCCTGCAAAAGTAACCTGGATGGCTT 3655
Db 2820 TGCCAGCTTGGGCGCCCTCTGTTAAGTTGGGAAGCCCTGCAAAAGTAACCTGGATGGCTT 2879
Qy 3656 TCTTGCCGCCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGGATGAG 3715
Db 2880 TCTTGCCGCCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGGATGAG 2939
Qy 3716 GATGTTTTCCGATGATTAAGCAAGATGATTTGCAAGCAGGTTCTCCGCGCCTTGGGTGG 3775
Db 2940 GATGTTTTCCGATGATTAAGCAAGATGATTTGCAAGCAGGTTCTCCGCGCCTTGGGTGG 2999
Qy 3776 AGAGGCTATTCGGCTATGACTGGGCAACACAGACAATCGGCTGCTCTGATGCCCGCGTGT 3835
Db 3000 AGAGGCTATTCGGCTATGACTGGGCAACACAGACAATCGGCTGCTCTGATGCCCGCGTGT 3059
Qy 3836 TCCGCTCTCAGCGCAGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCC 3895
Db 3060 TCCGCTCTCAGCGCAGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCC 3119

Qy 3896 TGAATGAATGCAAGGACGAGCAGCGGGCTATCGTGGCTGGCCACGACGGCGGTTCCTT 3955
Db 3120 TGAATGAATGCAAGGACGAGCAGCGGGCTATCGTGGCTGGCCACGACGGCGGTTCCTT 3179
Qy 3956 GCGCAGCTGTCTCGACGTTGTCTCACTGAAGCGGGAAGGACTGTGCTGTATTTGGCGAAG 4015
Db 3180 GCGCAGCTGTCTCGACGTTGTCTCACTGAAGCGGGAAGGACTGTGCTGTATTTGGCGAAG 3239
Qy 4016 TGCCGGGGCAGGATCTCTGTCACTCATCTCACCTTGTCTCTGCCGAGAAAGTATCCATCATGG 4075
Db 3240 TGCCGGGGCAGGATCTCTGTCACTCATCTCACCTTGTCTCTGCCGAGAAAGTATCCATCATGG 3299
Qy 4076 CTGATGCAATGCGCGGCTGTCACTGTGATCCGGCTTACCTGCGCCATTCACCAACCAAG 4135
Db 3300 CTGATGCAATGCGCGGCTGTCACTGTGATCCGGCTTACCTGCGCCATTCACCAACCAAG 3359
Qy 4136 CGAAACATCGCATCGAGCGAGCAGTACTCGATGGAAGCCGGTCTTGTGATCAGGATG 4195
Db 3360 CGAAACATCGCATCGAGCGAGCAGTACTCGATGGAAGCCGGTCTTGTGATCAGGATG 3419
Qy 4196 ATCTGGACGAGAGCATCAGGGGCTCGCGCAGCGCAACTGTTCCGCAAGGCTCAAGGGCC 4255
Db 3420 ATCTGGACGAGAGCATCAGGGGCTCGCGCAGCGCAACTGTTCCGCAAGGCTCAAGGGCC 3479
Qy 4256 GCATGCCCGCAGCGGAGGATCTCGTCTGTGACCCATGCGCATGCTGTGCGCAATATCA 4315
Db 3480 GCATGCCCGCAGCGGAGGATCTCGTCTGTGACCCATGCGCATGCTGTGCGCAATATCA 3539
Qy 4316 TGGTGGAAATGCGCGCTTTTCTGGATTCATTCGATGTGCGCGGCTGGTGTGCGGACC 4375
Db 3540 TGGTGGAAATGCGCGCTTTTCTGGATTCATTCGATGTGCGCGGCTGGTGTGCGGACC 3599
Qy 4376 GCTATCAGGACATAGCTTGGCTTACCCGTGATATTTGCTGAGAGCTTGGCGCGAATGGG 4435
Db 3600 GCTATCAGGACATAGCTTGGCTTACCCGTGATATTTGCTGAGAGCTTGGCGCGAATGGG 3659
Qy 4436 CTGACCGCTTCTCTGCTTTTACGGTATCGCGCTCCCGATTCGACGCGCATCGCTTCT 4495
Db 3660 CTGACCGCTTCTCTGCTTTTACGGTATCGCGCTCCCGATTCGACGCGCATCGCTTCT 3719
Qy 4496 ATCGCTTCTTGACGAGTTCTTCTGACGCGGACTCTCGGGTTTCGATA 4542
Db 3720 ATCGCTTCTTGACGAGTTCTTCTGACGCGGACTCTCGGGTTTCGATA 3766

RESULT 12

AAx90484
ID AAx90484 standard; DNA; 5874 BP.
AC AAx90484;
XX
XX 29-SEP-1999 (first entry)
DT
XX
DE Plasmid retroviral vector pLXSN nucleotide sequence.
XX
XX Plasmid retroviral vector; expression system; immunogenic; gene therapy;
KW immune response; immunosuppression; gene delivery; therapeutic;
KW MHC-I autoimmune disease; tumour; ss.
XX
OS Synthetic.
XX
XX WO9936562-A1.
XX
XX 22-JUL-1999.
PD
XX
XX 13-JAN-1999; 99WO-US000733.
PF
XX
XX 14-JAN-1998; 98US-0071409P.
PR
XX
XX (HUMA-) HUMAN GENE THERAPY RES INST.
PA
XX
XX Radosevich TJ, Link CJ;

XX 22-JUL-1999.
XX 13-JAN-1999; 99WO-US000733.
XX 14-JAN-1998; 98US-0071409P.
XX (HUMA-) HUMAN GENE THERAPY RES INST.
XX Radosevich TJ, Link CJ;
XX WPI; 1999-468988/39.
XX Expression system containing therapeutic gene and an immunosuppressor
XX gene useful for treating an MHC-I autoimmune disease or killing tumor
XX cells.
XX Claim 25; Page 134-137; 154pp; English.
XX The present invention describes a nucleotide expression system for the
XX introduction of a therapeutic gene comprising: (i) a nucleotide sequence
XX encoding an immune suppression gene; (ii) a promoter; and (iii) a
XX transcription termination signal, where the system is able to inhibit,
XX evade or eliminate a recipient cell immune response to the therapeutic
XX gene when the gene is transformed into a recipient cell. AAX90481 to
XX AAX90484 represent specifically claimed plasmid retroviral vector
XX nucleotide sequences from the present invention. The expression system
XX and vectors containing it can be used for gene therapy, for treating an
XX MHC-I autoimmune disease or for killing tumour cells. The expression
XX system contains an immunosuppressive gene which prevents host rejection
XX of the vector
XX
XX Sequence 6141 BP; 1404 A; 1698 C; 1603 G; 1436 T; 0 U; 0 Other;
XX
XX Query Match 28.7%; Score 1486; DB 2; Length 6141;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 3692 GATCTGATCAAGACAGGATGAGGATCGTTCCGATGATTAACAGATGGATGACG 3751
DB 2298 GATCTGATCAAGACAGGATGAGGATCGTTCCGATGATTAACAGATGGATGACG 2357
QY 3752 CAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCCGGCTATGACTGGGCAACAACAGACAA 3811
DB 2358 CAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCCGGCTATGACTGGGCAACAACAGACAA 2417
QY 3812 TCGGCTGCTCTGATGCGCCGCTGTTCGGGCTGTGACGCGAGGGCGCCGGTCTCTTTTG 3871
DB 2418 TCGGCTGCTCTGATGCGCCGCTGTTCGGGCTGTGACGCGAGGGCGCCGGTCTCTTTTG 2477
QY 3872 TCAAGACCGACTGTCCGGTGCCTTGATGAATGAATGACGAGGAGGCGCGGCTATCGT 3931
DB 2478 TCAAGACCGACTGTCCGGTGCCTTGATGAATGACGAGGAGGCGCGGCTATCGT 2537
QY 3932 GCTGGCCACAGCGGGCGTTCCTTGGCAGCTGTGCTCGAGCTGTGCTCACTCAAGCGGAA 3991
DB 2538 GCTGGCCACAGCGGGCGTTCCTTGGCAGCTGTGCTCGAGCTGTGCTCACTCAAGCGGAA 2597
QY 3992 GGGACTGGCTGTATTGGGCGAAGTGC CGGGGCGAGGATCTCTGTCACTCTCACCTTGCTC 4051
DB 2598 GGGACTGGCTGTATTGGGCGAAGTGC CGGGGCGAGGATCTCTGTCACTCTCACCTTGCTC 2657
QY 4052 CTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGGCGGCTGATACGCTTATCCGG 4111
DB 2658 CTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGGCGGCTGATACGCTTATCCGG 2717
QY 4112 CTACTCGCCATTTCGACCACCAAGCGAACAATCGATCGAGCGAGCTACTCGGATGG 4171
DB 2718 CTACTCGCCATTTCGACCACCAAGCGAACAATCGATCGAGCGAGCTACTCGGATGG 2777
QY 4172 AAGCGGCTCTTGTGATCAGGATGATCTGGACGAGAGCATCAGGGGCTCGCGCCAGCGC 4231
DB 2778 AAGCGGCTCTTGTGATCAGGATGATCTGGACGAGAGCATCAGGGGCTCGCGCCAGCGC 2837

QY 4232 AACTGTTCCGCGAGCTCAAGCGCGCATGCCGACGCGGAGGATCTCGTCTGACCCCATG 4291
DB 2838 AACTGTTCCGCGAGCTCAAGCGCGCATGCCGACGCGGAGGATCTCGTCTGACCCCATG 2897
QY 4292 GCGATGCTCTGTTGCCGAATATCATGTGGAAATATGCGCGCTTTTCTGGATTCATGACT 4351
DB 2898 GCGATGCTCTGTTGCCGAATATCATGTGGAAATATGCGCGCTTTTCTGGATTCATGACT 2957
QY 4352 GTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTCGCTACCGCTGATATTG 4411
DB 2958 GTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTCGCTACCGCTGATATTG 3017
QY 4412 CTGAAGAGCTTGGCGGGAATGGCTGACCGCTTCTCTGCTTTTACGGTATCCCGCTC 4471
DB 3018 CTGAAGAGCTTGGCGGGAATGGCTGACCGCTTCTCTGCTTTTACGGTATCCCGCTC 3077
QY 4472 CCGATTCCGACGCGCATCGCTTCTATCGCTTCTTACGAGTCTTCTTGAGCGGGACTCT 4531
DB 3078 CCGATTCCGACGCGCATCGCTTCTATCGCTTCTTACGAGTCTTCTTGAGCGGGACTCT 3137
QY 4532 GGGGTTCCGATAAAATAAAGATTTTATTAGTCTCCAGAAAAAGGGGGGAATGAAAGACC 4591
DB 3138 GGGGTTCCGATAAAATAAAGATTTTATTAGTCTCCAGAAAAAGGGGGGAATGAAAGACC 3197
QY 4592 CCACCTGTAGGTTGGCAAGCTAGCTTAAAGTAAGCGCATTTTGCAGGSCATGGAAAAATA 4651
DB 3198 CCACCTGTAGGTTGGCAAGCTAGCTTAAAGTAAGCGCATTTTGCAGGSCATGGAAAAATA 3257
QY 4652 CATAACTGAGATAGAGAAGTTTCAGATCAAGGTCAGGAACAGATGGAACAGTGAATATG 4711
DB 3258 CATAACTGAGATAGAGAAGTTTCAGATCAAGGTCAGGAACAGATGGAACAGTGAATATG 3317
QY 4712 GGCCAAACAGGATATCTGTGTGAAGCAGTTCTGCCCCCGGCTCAGGGCCAAAGAACAGATG 4771
DB 3318 GGCCAAACAGGATATCTGTGTGAAGCAGTTCTGCCCCCGGCTCAGGGCCAAAGAACAGATG 3377
QY 4772 GAACAGCTGAATATGGGCCAAACAGGATATCTGTGTGAAGCAGTTCTGCCCCCGGCTCAG 4831
DB 3378 GAACAGCTGAATATGGGCCAAACAGGATATCTGTGTGAAGCAGTTCTGCCCCCGGCTCAG 3437
QY 4832 GGCCAAACAGGATATCTGTGTGAAGCAGTTCTGCCCCCGGCTCAGGGCCAAAGAACAGATG 4891
DB 3438 GGCCAAACAGGATATCTGTGTGAAGCAGTTCTGCCCCCGGCTCAGGGCCAAAGAACAGATG 3497
QY 4892 AGATGTTTCCAGGGTGCCCCCAAGGACCTGAAATGACCTGTGCTTATTTGAACTAACCA 4951
DB 3498 AGATGTTTCCAGGGTGCCCCCAAGGACCTGAAATGACCTGTGCTTATTTGAACTAACCA 3557
QY 4952 ATCAGTTTCGCTTCTCGCTTCTGTTCGCGCGCTTCTGCTCCCCGAGCTCAATAAAGAGCC 5011
DB 3558 ATCAGTTTCGCTTCTCGCTTCTGTTCGCGCGCTTCTGCTCCCCGAGCTCAATAAAGAGCC 3617
QY 5012 CACAAACCCCTCACTCGGGGGCCAGATCCTCGATTGACTGAGTCCGCGGGTACCCGCTGT 5071
DB 3618 CACAAACCCCTCACTCGGGGGCCAGATCCTCGATTGACTGAGTCCGCGGGTACCCGCTGT 3677
QY 5072 ATCAATAAACCCTTTTCAGATTGTCATCCGACTTGTGGTCTCTGCTGTTCCTTGGGAGGGT 5131
DB 3678 ATCAATAAACCCTTTTCAGATTGTCATCCGACTTGTGGTCTCTGCTGTTCCTTGGGAGGGT 3737
QY 5132 CTCTCTGAGTGATGATGACTACCGCTCAGCGGGGCTCTTTCAATTTGG 5177
DB 3738 CTCTCTGAGTGATGATGACTACCGCTCAGCGGGGCTCTTTCAATTTGG 3783
RESULT 14
AAX90482
ID AAX90482 standard; DNA; 6522 BP.
XX
XX AAX90482;
XX AC
XX AC
DT 29-SEP-1999 (first entry)

XX Plasmid retroviral vector pLUSN nucleotide sequence.
XX
XX Plasmid retroviral vector; expression system; immunogenic; gene therapy;
KW immune response; immunosuppression; gene delivery; therapeutic;
KW MHC-I autoimmune disease; tumour; ss.
XX
XX Synthetic.
XX
XX WO936562-A1.
XX
XX 22-JUL-1999.
XX
XX 13-JAN-1999; 99WO-US000733.
XX
XX 14-JAN-1998; 98US-0071409P.
XX
XX (HUMA-) HUMAN GENE THERAPY RES INST.
XX
XX Radosevich TV, Link CJ;
XX
XX WPI; 1999-468988/39.
XX
XX Expression system containing therapeutic gene and an immunosuppressor
PT gene useful for treating an MHC-I autoimmune disease or killing tumor
PT cells.
XX
XX Claim 25; Page 137-140; 15app; English.
XX
XX The present invention describes a nucleotide expression system for the
CC introduction of a therapeutic gene comprising: (i) a nucleotide sequence
CC encoding an immune suppression gene; (ii) a promoter; and (iii) a
CC transcription termination signal, where the system is able to inhibit,
CC evade or eliminate a recipient cell immune response to the therapeutic
CC gene when the gene is transformed into a recipient cell. AAX90481 to
CC AAX90484 represent specifically claimed plasmid retroviral vector
CC nucleotide sequences from the present invention. The expression system
CC and vectors containing it can be used for gene therapy, for treating an
CC MHC-I autoimmune disease or for killing tumour cells. The expression
CC system contains an immunosuppressive gene which prevents host rejection
CC of the vector
XX
XX Sequence 6522 BP; 1469 A; 1792 C; 1704 G; 1557 T; 0 U; 0 Other;
SQ
Query Match 28.7%; Score 1486; DB 2; Length 6522;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3692 GATCTGATCAAGAGACAGAGATGAGATCGTTTCGATGATTGAACAGATGGATTGACG 3751
Db 2679 GATCTGATCAAGAGACAGAGATGAGATCGTTTCGATGATTGAACAGATGGATTGACG 2738
Qy 3752 CAGGTTCTCCGGCGCGCTTGGGTGAGAGGCTATTTCGGCTATGACTGGGCACACAGACAA 3811
Db 2739 CAGGTTCTCCGGCGCGCTTGGGTGAGAGGCTATTTCGGCTATGACTGGGCACACAGACAA 2798
Qy 3812 TCGGCTGCTCTGATGCCCGCGCTTTCGGCTGTGACGCGAGGCGCGCGGTTCTTTTG 3871
Db 2799 TCGGCTGCTCTGATGCCCGCGCTTTCGGCTGTGACGCGAGGCGCGCGGTTCTTTTG 2858
Qy 3872 TCAAGACCGA CTTGTCGGTCCCTGTAATGAATGACGAGGACGAGCGCGGCTATCGT 3931
Db 2859 TCAAGACCGA CTTGTCGGTCCCTGTAATGAATGACGAGGACGAGCGCGGCTATCGT 2918
Qy 3932 GGCTGGCCACACCGGCGCTTCTTGGCGAGCTGTGCTCGAGCTTGTCACTGAACGGGAA 3991
Db 2919 GGCTGGCCACACCGGCGCTTCTTGGCGAGCTGTGCTCGAGCTTGTCACTGAACGGGAA 2978
Qy 3992 GGGACTGACTGCTATTGGGCGAAGTGCGGCGGAGGATCTCCTGTCTCATCTCACCTTGTCTC 4051
Db 2979 GGGACTGACTGCTATTGGGCGAAGTGCGGCGGAGGATCTCCTGTCTCATCTCACCTTGTCTC 3038
Qy 4052 CTGCCGAGAAAGTATCCATCATGCTGATGCAATGCGCGGCTGCATACGCTTGTACCGG 4111

Db 3039 CTGCCGAGAAAGTATCCATCATGCTGATGCAATGCGCGGCTGCATACGCTTATCCGG 3098
Qy 4112 CTACTCTCCCATTCGACACCAAGCAAAACATCCATCGATCGAGCGACGACTACTCGGATGG 4171
Db 3099 CTACTCTCCCATTCGACACCAAGCAAAACATCCATCGATCGAGCGAGCAGTACTCGGATGG 3158
Qy 4172 AAGCCGCTCTTGTGCTGATCAGGATGATCTGACGAAAGAGCATCAGGGGCTCGCGCAGCG 4231
Db 3159 AAGCCGCTCTTGTGCTGATCAGGATGATCTGACGAAAGAGCATCAGGGGCTCGCGCAGCG 3218
Qy 4232 AACTGTTGCGCAGGCTCAAGCGCGCATGCGCGAGGAGGATCTGCTGCGTACCCCATG 4291
Db 3219 AACTGTTGCGCAGGCTCAAGCGCGCATGCGCGAGGAGGATCTGCTGCGTACCCCATG 3278
Qy 4292 GCATGCTCTGCTTCCGCAATATCATGCTGGAATAATGCGCGCTTTTCTGGATTCACTCACT 4351
Db 3279 GCATGCTCTGCTTCCGCAATATCATGCTGGAATAATGCGCGCTTTTCTGGATTCACTCACT 3338
Qy 4352 GTGGCCGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGCTGATATG 4411
Db 3339 GTGGCCGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGCTGATATG 3398
Qy 4412 CTGAGAGCTTGGCGGCAATGGGCTGACCGCTTCTCTGCTTCTACGGTATCGCGCTC 4471
Db 3399 CTGAGAGCTTGGCGGCAATGGGCTGACCGCTTCTCTGCTTCTACGGTATCGCGCTC 3458
Qy 4472 CCGATTCCGAGCGATCGCTTCTATCGCTTCTTGTAGAGTCTTCTTGACGGGACTCT 4531
Db 3459 CCGATTCCGAGCGATCGCTTCTATCGCTTCTTGTAGAGTCTTCTTGACGGGACTCT 3518
Qy 4532 GGGGTTTCGATAAAATAAAAGATTTTATTAGTCTCCAGAAAAAGGGGGAATGAAAGACC 4591
Db 3519 GGGGTTTCGATAAAATAAAAGATTTTATTAGTCTCCAGAAAAAGGGGGAATGAAAGACC 3578
Qy 4592 CCACCTGTAGGTTGGCAAGCTAGCTTAACTAGCCCATTTTGAAGGCATGGAAGAAATA 4651
Db 3579 CCACCTGTAGGTTGGCAAGCTAGCTTAACTAGCCCATTTTGAAGGCATGGAAGAAATA 3638
Qy 4652 CATAACTGAGAAATAGAGAGTTTCAGATCAAGGTCAGGAAACAGATGGAACAGCTCAATATG 4711
Db 3639 CATAACTGAGAAATAGAGAGTTTCAGATCAAGGTCAGGAAACAGATGGAACAGCTCAATATG 3698
Qy 4712 GGCCAAACAGGATATCTGTGTAGAGTTCCTGCCCCGGCTCAGGGCCAGAAACAGATG 4771
Db 3699 GGCCAAACAGGATATCTGTGTAGAGTTCCTGCCCCGGCTCAGGGCCAGAAACAGATG 3758
Qy 4772 GAACAGCTGAATATGGGCCAAAACAGGATATCTGTGTAGAGTTCCTGCCCCGGCTCAG 4831
Db 3759 GAACAGCTGAATATGGGCCAAAACAGGATATCTGTGTAGAGTTCCTGCCCCGGCTCAG 3818
Qy 4832 GGCCAAACAGAGATGTTCCCGAGATGCGGTCCAGCCCTCAGCAGTTCCTAGAGAACCATC 4891
Db 3819 GGCCAAACAGAGATGTTCCCGAGATGCGGTCCAGCCCTCAGCAGTTCCTAGAGAACCATC 3878
Qy 4892 AGATGTTTCCAGGTCGCCCAAGGACCTGAAATGACCTGTGCTTATTTGAACTAACCA 4951
Db 3879 AGATGTTTCCAGGTCGCCCAAGGACCTGAAATGACCTGTGCTTATTTGAACTAACCA 3938
Qy 4952 ATCAGTTTCGCTTCTCGCTTCTGTGTCGCGCTTCTGCTCCCGAGCTCAATAAAGAGCC 5011
Db 3939 ATCAGTTTCGCTTCTCGCTTCTGTGTCGCGCTTCTGCTCCCGAGCTCAATAAAGAGCC 3998
Qy 5012 CACAACCCCTCCTCAGGGGCGCAGTCTCCGATGACTGAGTCCCGGCTACCCGCTGT 5071
Db 3999 CACAACCCCTCCTCAGGGGCGCAGTCTCCGATGACTGAGTCCCGGCTACCCGCTGT 4058
Qy 5072 ATCCAAATAAACCCCTCTTTCGAGTTGCACTCCGACTTGTGCTCTCGCTTCTTGGAGGGT 5131
Db 4059 ATCCAAATAAACCCCTCTTTCGAGTTGCACTCCGACTTGTGCTCTCGCTTCTTGGAGGGT 4118
Qy 5132 CTCTCTGAGTGAATGATACCCGCTCAGCGGGGCTTTTTCATTTGG 5177

Db 4119 CTCCTCTGAGTATTGACTACCGGTACGGGGGTCTTTTCATTGG 4164

RESULT 15

AD82653
ID AD82653 standard; DNA; 6620 BP.

AC AD82653;

DT 29-JAN-2004 (first entry)

XX Plasmid vector pUESN nucleotide sequence SEQ ID NO:3.

XX recombinant DNA construct;
KW humanised red shifted green fluorescent protein; HRGFP; fluorescence;
KW gene transfer; gene expression; gene therapy; bone marrow processing;
KW plasmid vector; gene; db.

XX Synthetic.

XX US2002015979-A1.

XX 07-FEB-2002.

XX 21-JAN-1997; 97US-00786531.

XX 22-JAN-1996; 96US-0010371P.

XX (LINK/) LINK C J.

PA (LEVY/) LEVY J P.

PA (WANG/) WANG S.

PA (SERE/) SEREGINA T.

XX Link C J, Levy J P, Wang S, SerEGINA T;

XX WPI; 2002-239224/29.

XX Recombinant DNA construct for identifying transformed cells to allow for
PT direct observation of transferred genes into living cells, comprising
PT nucleotide sequence encoding red shifted green fluorescent protein.

PS Claim 17; SEQ ID NO 3; 95pp: English.

XX The present invention describes a recombinant DNA construct (I),
CC comprising a nucleotide sequence encoding upon expression a humanised red
CC shifted green fluorescent protein (HRGFP), and a promoter sequence and a
CC termination sequence functionally coupled to the coding sequence. Also
CC described: (1) a plasmid vector DNA sequence (II) enabling replication of
CC the vector or in a host cell and (I); (2) infectious virus (III)
CC comprising retroviral RNA transcribed from (I) in a host cell capable of
CC viral packaging; and (3) a mammalian cell (IV) or a cell derived from it,
CC comprising at least one copy of (I). (I) is useful for identifying
CC transformed cells to allow for direct observation of transferred genes
CC into living cells, by introducing (I) into the cell and measuring
CC fluorescence of HRGFP, expressed by the transformed cells. The
CC fluorescence is measured by the absorbance at 490-760 nm, and the
CC excitation is from 420-470 nm. (I) is also useful for selecting
CC lymphocytes which have been transfected with a vector, by transforming
CC cells with (I) and sorting the cells with a fluorescence activated cell
CC sorter (FACS). The expression of the HRGFP gene in living eukaryotic
CC cells will advance the study of gene transfer, gene expression and gene
CC product function in vitro or in vivo, particularly for human gene therapy
CC applications. (I) is useful in bone marrow processing and other
CC applications requiring FACS analyses of living cells. The present
CC sequence represents a specifically claimed plasmid vector nucleotide
CC sequence, which is used in the exemplification of the present invention.
CC N.B. The SEQ ID NO:8 in the Sequence Listing do not correspond with SEQ
CC ID NO:8 mentioned in the specification: where SEQ ID NO:1 to 6 of the
CC Sequence Listing represent the sequences in figures 22, 17, 18, 20, 19
CC and 21 respectively; SEQ ID NO:7 to 10 of the Sequence Listing represents
CC SEQ ID NO:2 to 5 on page 7 of example 1; and SEQ ID NO:1 of the
CC specification given in figure 7 is not given in the Sequence Listing.
XX

SQ Sequence 6620 BP; 1527 A; 1863 C; 1728 G; 1502 T; 0 U; 0 Other;
Query Match 28.7%; Score 1486; DB 6; Length 6620;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3692 GATCTGATCAAGACAGACAGGATGAGGATCGTTTCCGATGATTGAACAAGATTGCACG 3751
Db 2777 GATCTGATCAAGACAGACAGGATGAGGATCGTTTCCGATGATTGAACAAGATTGCACG 2836
Qy 3752 CAGGTTCTCCGGCCGCTTGGGTGAGAGGCTATTTCGGCTATGACTGGGCAACAAGACAA 3811
Db 2837 CAGGTTCTCCGGCCGCTTGGGTGAGAGGCTATTTCGGCTATGACTGGGCAACAAGACAA 2896
Qy 3812 TCGGCTGCTGTGATGCCCGCTGTTCCGGCTGTGACGCGAGGGCGCCCGGTTCTTTTGG 3871
Db 2897 TCGGCTGCTGTGATGCCCGCTGTTCCGGCTGTGACGCGAGGGCGCCCGGTTCTTTTGG 2956
Qy 3872 TCAAGACCGACCTGTTCGGGTGCCCTGAATGAATGCAGAGCAGGCGAGCGCGGTATCGT 3931
Db 2957 TCAAGACCGACCTGTTCGGGTGCCCTGAATGAATGCAGAGCAGGCGAGCGCGGTATCGT 3016
Qy 3932 GGCTGCGCACGACGGGCGTTTCTTGGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAA 3991
Db 3017 GGCTGCGCACGACGGGCGTTTCTTGGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAA 3076
Qy 3992 GGGACTGGCTGCTATTGGGGCGAAGTCCGGGGCAGGATCTCTGTGTCATCTCACCTTGCTC 4051
Db 3077 GGGACTGGCTGCTATTGGGGCGAAGTCCGGGGCAGGATCTCTGTGTCATCTCACCTTGCTC 3136
Qy 4052 CTGCCGAGAAAGTATTCATCATGCTGATGCAATGCGGGCGGCTGCATACGCTTCGATCCGG 4111
Db 3137 CTGCCGAGAAAGTATTCATCATGCTGATGCAATGCGGGCGGCTGCATACGCTTCGATCCGG 3196
Qy 4112 CTACCTGCCATTGACCAACAAGCGAAACATGCAATCGACGCGAGCAGTACTCGGATGG 4171
Db 3197 CTACCTGCCATTGACCAACAAGCGAAACATGCAATCGACGCGAGCAGTACTCGGATGG 3256
Qy 4172 AAGCGGTCTTGTGATCAGGATGATCTGACGAGAGGATCAGGGGCTCGCGCCAGCG 4231
Db 3257 AAGCGGTCTTGTGATCAGGATGATCTGACGAGAGGATCAGGGGCTCGCGCCAGCG 3316
Qy 4232 AACTGTTTGGCAGGCTCAAGCGCGCATGCGCCGAGGAGTCTCGTGTGACCCCATG 4291
Db 3317 AACTGTTTGGCAGGCTCAAGCGCGCATGCGCCGAGGAGTCTCGTGTGACCCCATG 3376
Qy 4292 GCGATGCTCTGTCGCCGAATATCATGTGGAAATATGCGCCGCTTTTCTGGATTTCAGACT 4351
Db 3377 GCGATGCTCTGTCGCCGAATATCATGTGGAAATATGCGCCGCTTTTCTGGATTTCAGACT 3436
Qy 4352 GTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGGTGATATTG 4411
Db 3437 GTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGGTGATATTG 3496
Qy 4412 CTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCTCGTGTGCTTTACGGTATCCCGCTC 4471
Db 3497 CTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCTCGTGTGCTTTACGGTATCCCGCTC 3556
Qy 4472 CCGATTCCGACGCGATCGCTTCTTATCGCTTCTTGTGACGAGTCTTCTGACGCGGACTCT 4531
Db 3557 CCGATTCCGACGCGATCGCTTCTTATCGCTTCTTGTGACGAGTCTTCTGACGCGGACTCT 3616
Qy 4532 GGGGTTTCGATAAATAAAGATTTTATTAGTCTCCAGAAAGGGGGGGAATGAAGACC 4591
Db 3617 GGGGTTTCGATAAATAAAGATTTTATTAGTCTCCAGAAAGGGGGGGAATGAAGACC 3676
Qy 4592 CCACCTCTAGGTTTGGCAAGCTAGCTTAAGTAAGCCCATTTTTCGAAGGCATGGAATAATA 4651
Db 3677 CCACCTCTAGGTTTGGCAAGCTAGCTTAAGTAAGCCCATTTTTCGAAGGCATGGAATAATA 3736
Qy 4652 CATAACTGAGAATAGAGAAGTTCAGATCAAGGTCAGGAACAGATGGAACAGCTGAATATG 4711
Db 3737 CATAACTGAGAATAGAGAAGTTCAGATCAAGGTCAGGAACAGATGGAACAGCTGAATATG 3796

QY 4712 GGCCAAACAGGATATCTGTGTAAGCAGTTCTGCCCCGGCTCAGGGGCCAAGAACAGATG 4771
Db |||||
QY 3797 GGCCAAACAGGATATCTGTGTAAGCAGTTCTGCCCCGGCTCAGGGGCCAAGAACAGATG 3856
Db |||||
QY 4772 GAAACAGCTGATATGGCCCAACAGGATATCTGTGTAAGCAGTTCTGCCCCGGCTCAG 4831
Db |||||
QY 3857 GAAACAGCTGATATGGCCCAACAGGATATCTGTGTAAGCAGTTCTGCCCCGGCTCAG 3916
Db |||||
QY 4832 GGCCAAAGAACAGATGGTCCCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGAGAACCATC 4891
Db |||||
QY 3917 GGCCAAAGAACAGATGGTCCCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGAGAACCATC 3976
Db |||||
QY 4892 AGATGTTTCCAGGGTGCCCCAAGAACCTGAAATGACCCCTGTGCTTATTTGAACCTAACCA 4951
Db |||||
QY 3977 AGATGTTTCCAGGGTGCCCCAAGAACCTGAAATGACCCCTGTGCTTATTTGAACCTAACCA 4036
Db |||||
QY 4952 ATCAGTTCGCTTCTCGCTTCTGTTCCGCGCTTCTGCTCCCGAGCTCAATAAAGAGCC 5011
Db |||||
QY 4037 ATCAGTTCGCTTCTCGCTTCTGTTCCGCGCTTCTGCTCCCGAGCTCAATAAAGAGCC 4096
Db |||||
QY 5012 CACAACCCCTCACTCGGGGGCCAGTCTCTCGATTGACTGAGTCGCCGGGTACCCGTGT 5071
Db |||||
QY 4097 CACAACCCCTCACTCGGGGGCCAGTCTCTCGATTGACTGAGTCGCCGGGTACCCGTGT 4156
Db |||||
QY 5072 ATCCAATAAACCCCTCTTGCACTTGCATCCGACTTGTGCTCTGCTGTTCTTGGAGGGT 5131
Db |||||
QY 4157 ATCCAATAAACCCCTCTTGCACTTGCATCCGACTTGTGCTCTGCTGTTCTTGGAGGGT 4216
Db |||||
QY 5132 CTCCTCTGAGTGATTGACTACCCGTACGGGGGTCTTTCAATTGG 5177
Db |||||
QY 4217 CTCCTCTGAGTGATTGACTACCCGTACGGGGGTCTTTCAATTGG 4262
Db |||||

Search completed: February 27, 2006, 09:24:01
Job time : 2890 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
 Copyright (c) 1993 - 2006 Bioceleration Ltd.
 OM nucleic - nucleic search, using sw model
 Run on: February 27, 2006, 08:36:14 ; Search time 3728 Seconds
 (without alignments)
 11483.522 Million cell updates/sec

Title: US-10-618-570-1
 Perfect score: 5177
 Sequence: 1 ttgaagaccaccaccgta.....agegggggtttttcattgg 5177

Scoring table: IDENTITY_NUC
 Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Published Applications NA Main.*
 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5177	100.0	5177	8	US-10-618-570-1
2	1588.6	30.7	5130	3	US-09-897-006-9
3	1588.6	30.7	5130	3	US-09-897-511A-9
4	1588.6	30.7	5130	6	US-10-397-079-9
5	1588.6	30.7	5130	8	US-10-759-315-9
6	1588.6	30.7	5130	9	US-10-947-881-9
7	1588.6	30.7	5130	10	US-11-018-895-9
8	1588.6	30.7	5130	10	US-11-036-557-9
9	1486	28.7	6620	2	US-08-786-531B-3
10	1481	28.6	6046	6	US-10-164-965A-1
11	1463	28.3	6825	9	US-10-987-388-36
12	1463	28.3	6851	9	US-10-987-388-35
13	1455.2	28.1	5856	7	US-10-359-397-1
14	1455.2	28.1	5865	5	US-10-098-035-3
15	1448	28.0	6444	3	US-09-808-743-2
16	1434	27.7	6700	3	US-09-759-152-1
17	1434	27.7	8518	3	US-09-759-152-4
18	1425	27.5	5689	8	US-10-836-806-1
19	1174	22.7	7617	7	US-10-401-000-1
20	1170.4	22.6	3974	3	US-09-027-287-50
21	1170.4	22.6	3974	3	US-09-252-656B-50
22	1170.4	22.6	3974	3	US-09-972-912-8
23	1170.4	22.6	3974	3	US-09-902-684-14

Sequence 16, Appl
 Sequence 25, Appl
 Sequence 147, Appl
 Sequence 12, Appl
 Sequence 16, Appl
 Sequence 12, Appl
 Sequence 33, Appl
 Sequence 7, Appl
 Sequence 4, Appl
 Sequence 9, Appl
 Sequence 16, Appl
 Sequence 16, Appl
 Sequence 16, Appl
 Sequence 147, Appl
 Sequence 147, Appl
 Sequence 4, Appl
 Sequence 9, Appl
 Sequence 16, Appl
 Sequence 16, Appl
 Sequence 16, Appl
 Sequence 147, Appl
 Sequence 147, Appl
 Sequence 50, Appl
 Sequence 25, Appl
 Sequence 14, Appl
 Sequence 16, Appl
 Sequence 147, Appl
 Sequence 4, Appl

ALIGNMENTS

RESULT 1
 US-10-618-570-1
 ; Sequence 1, Application US/10618570
 ; Publication No. US20040185059A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Yla-Herttuala, Seppo
 ; APPLICANT: Kulomaa, Markku
 ; APPLICANT: Lehtolainen, Pauliina
 ; APPLICANT: Marjomaki, Varpu
 ; APPLICANT: Airenne, Kari
 ; TITLE OF INVENTION: Biotin-Binding Receptor Molecules
 ; FILE REFERENCE: GJE-48
 ; CURRENT APPLICATION NUMBER: US/10/618,570
 ; CURRENT FILING DATE: 2003-07-11
 ; PRIOR APPLICATION NUMBER: US/09/622,804
 ; PRIOR FILING DATE: 2002-12-10
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 5177
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Recombinant cDNA encoding fusion protein
 ; NAME/KEY: CDS
 ; LOCATION: (1071)..(2270)
 ; OTHER INFORMATION:
 US-10-618-570-1

Query Match 100.0%; Score 5177; DB 8; Length 5177;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 5177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TTTGAAAGACCCCAACCGGTAGTGGCAAGCTTAGCTTAAGTAACGCCACTTTTGAAGGCAT 60
 Db 1 TTTGAAAGACCCCAACCGGTAGTGGCAAGCTTAGCTTAAGTAACGCCACTTTTGAAGGCAT 60
 Qy 61 GGAATAATACATACTAGTAATAGAAAAGTTTCAGATCAAGTCAAGTAAGAAAACAGC 120
 Db 61 GGAATAATACATACTAGTAATAGAAAAGTTTCAGATCAAGTCAAGTAAGAAAACAGC 120
 Qy 121 TGAATACCAACAGGATATCTGTGGTAAAGCGGTCTCTCCCGGCTCAGGCCCAAGACA 180
 Db 121 TGAATACCAACAGGATATCTGTGGTAAAGCGGTCTCTCCCGGCTCAGGCCCAAGACA 180

Qy	181	GATGAGACAGCTGAGTGATGGGCCAAACAGAGATATCTGTGTGTAAGCAGTTCCTGCGCCCGG	240
Db	181	GATGAGACAGCTGAGTGATGGGCCAAACAGAGATATCTGTGTGTAAGCAGTTCCTGCGCCCGG	240
Qy	241	CTCGGGGCCAAGACAGATGCTGCCAGATGCGGTCCAGCCCTCAGCAGTTCCTAGTGAA	300
Db	241	CTCGGGGCCAAGACAGATGCTGCCAGATGCGGTCCAGCCCTCAGCAGTTCCTAGTGAA	300
Qy	301	TCATCAGATGTTCTCAGGGTGCCCCAAAGGACCTGAAATGACCCCTGTACCTTATTTGCAAC	360
Db	301	TCATCAGATGTTCTCAGGGTGCCCCAAAGGACCTGAAATGACCCCTGTACCTTATTTGAC	360
Qy	361	TAAACAAATCAGTTTCGCTTCTCGCTTCTGTTTCGCGCGCTTCGCTCTCCGAGCTCAATAAA	420
Db	361	TAAACAAATCAGTTTCGCTTCTCGCTTCTGTTTCGCGCGCTTCGCTCTCCGAGCTCAATAAA	420
Qy	421	AGAGCCACAAACCCCTCACTCGGCGCGCCAGTCTTCCGATAGACTGCGTCCGCCGGGTAC	480
Db	421	AGAGCCACAAACCCCTCACTCGGCGCGCCAGTCTTCCGATAGACTGCGTCCGCCGGGTAC	480
Qy	481	CCGTATTCCCAATAAAGGCTCTTGCTGTTTGCATCCGAATCGTGCTCTCGCTGTTCTCTTG	540
Db	481	CCGTATTCCCAATAAAGGCTCTTGCTGTTTGCATCCGAATCGTGCTCTCGCTGTTCTCTTG	540
Qy	541	GGAGGGTCTCTCTGAGTGATTTGACTACCCACGACGGGGTCTTTCAATTTGGGGGCTCGT	600
Db	541	GGAGGGTCTCTCTGAGTGATTTGACTACCCACGACGGGGTCTTTCAATTTGGGGGCTCGT	600
Qy	601	CCGGGATTTGGAGACCCCTGCCAGGGACACCGACCCACCCGAGGAGTGAAGTCGCC	660
Db	601	CCGGGATTTGGAGACCCCTGCCAGGGACCAACCGACCCACCCGAGGAGTGAAGTCGCC	660
Qy	661	AGCAACTTATCTGTGTCTGTCGGAATGTCTAGTCTCTATGTTTCATGTTATGCGCCCTGCG	720
Db	661	AGCAACTTATCTGTGTCTGTCGGAATGTCTAGTCTCTATGTTTCATGTTATGCGCCCTGCG	720
Qy	721	TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGCGGACCCGCTGGTGGAATCGACGAGTT	780
Db	721	TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGCGGACCCGCTGGTGGAATCGACGAGTT	780
Qy	781	CTGAACACCCGGCCGCAACCCCTGGGAGACGTTCCAGGGAATTGTTGGGGGCGTTTTTGTGG	840
Db	781	CTGAACACCCGGCCGCAACCCCTGGGAGACGTTCCAGGGAATTGTTGGGGGCGTTTTTGTGG	840
Qy	841	CCCGACCTTGAGGAAGGAGTGCATGTGGAATCCGACCCCGTCAGGATATGTGGTCTTGCT	900
Db	841	CCCGACCTTGAGGAAGGAGTGCATGTGGAATCCGACCCCGTCAGGATATGTGGTCTTGCT	900
Qy	901	AGGAGACGAGAACCTTAAACAGTTCCTCGCTCCGCTCTGAAATTTTGTCTTCGGTTTGGAA	960
Db	901	AGGAGACGAGAACCTTAAACAGTTCCTCGCTCCGCTCTGAAATTTTGTCTTCGGTTTGGAA	960
Qy	961	CCGAAGCCGCGCTTGTCTGTGTCAGCCAAAGCTTGGGCTGCAGGTTCGACTTAGAGGA	1020
Db	961	CCGAAGCCGCGCTTGTCTGTGTCAGCCAAAGCTTGGGCTGCAGGTTCGACTTAGAGGA	1020
Qy	1021	TCAATTCCGGCACGAGTAAATCGGTGCTGCGCTCTTTAGGCACATATGAAGTATGSCACAGT	1080
Db	1021	TCAATTCCGGCACGAGTAAATCGGTGCTGCGCTCTTTAGGCACATATGAAGTATGSCACAGT	1080
Qy	1081	GGGATGACTTTTCTTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGGAAGTTCG	1140
Db	1081	GGGATGACTTTTCTTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGGAAGTTCG	1140
Qy	1141	ATGCTCGCTCAGTGACAGCTTTGCTTCTCCCAATCTTAAAAATGGGCCAACTCTTCAAG	1200
Db	1141	ATGCTCGCTCAGTGACAGCTTTGCTTCTCCCAATCTTAAAAATGGGCCAACTCTTCAAG	1200
Qy	1201	AGAGGATGAAGTCTTATAAAGCTGCATGATCACCCCTTATCTCATTTGTGTTGTAGTTC	1260
Db	1201	AGAGGATGAAGTCTTATAAAGCTGCATGATCACCCCTTATCTCATTTGTGTTGTAGTTC	1260
Qy	1261	TCGTGCCCATCATTTGGCATAGTGGCAGCTCAGCTCTCTGAAATGGGAAACGAGAATTTGCA	1320

1361	Db		TCGTGCCCATATTGGCATAGTGGCAGCTCAGTCTCTGTAAATGGGAAACGAAGAAATTGCA	1320
1321	Qy	CGGTTGCTCAGTTTAATGCAGATATATCTCCAAGTCCGGAAGCAAGGAAATGGCGATG	1380	
1321	Db	CGGTTGCTCAGTTTAATGCAGATATATCTCCAAGTCCGGAAGCAAGGAAATGGCGATG	1380	
1381	Qy	AAGATGAAATGAGATTTTCGAGAAAGCTGTGTAGTAAAGCGCATGAGCAACATGGAAGCAGAA	1440	
1381	Db	AAGATGAAATGAGATTTTCGAGAAAGCTGTGTAGTAAAGCGCATGAGCAACATGGAAGCAGAA	1440	
1441	Qy	TCAGTATCTTTTCAGATTAATGAAGCCAACTCTCTAGATGCTAAGAAATTTCCAAAATTTCA	1500	
1441	Db	TCCAGTATCTTTTCAGATTAATGAAGCCAACTCTCTAGATGCTAAGAAATTTCCAAAATTTCA	1500	
1501	Qy	GCATAACAACTGATCAAAGATTTTAATGATGCTCTTTTCAGCTTAAATTCCTTACTTTCTCT	1560	
1501	Db	GCATAACAACTGATCAAAGATTTTAATGATGCTCTTTTCAGCTTAAATTCCTTACTTTCTCT	1560	
1561	Qy	CCATCCAGGAACATGAGAAATATCATAGGGGATATCTCCAAGTCATTAGTAGGTCCTGAACA	1620	
1561	Db	CCATCCAGGAACATGAGAAATATCATAGGGGATATCTCCAAGTCATTAGTAGGTCCTGAACA	1620	
1621	Qy	CCACAGTACTTGATTTTCAGTTCAGTATTGAAACACTGAAATGGCAGAGTCCAAAGAAATG	1680	
1621	Db	CCACAGTACTTGATTTTCAGTTCAGTATTGAAACACTGAAATGGCAGAGTCCAAAGAAATG	1680	
1681	Qy	CATTTAACAAACAGAGGAGATCGGTAAATTTAGAGGAGCGTATATCAATGTCATCAGCAG	1740	
1681	Db	CATTTAACAAACAGAGGAGATCGGTAAATTTAGAGGAGCGTATATCAATGTCATCAGCAG	1740	
1741	Qy	AAATTAAGTCTCTAGATGAAAAACAAAGTATATTTGGAAACAGGAAATAAAAGGGGAAATGA	1800	
1741	Db	AAATTAAGTCTCTAGATGAAAAACAAAGTATATTTGGAAACAGGAAATAAAAGGGGAAATGA	1800	
1801	Qy	AACTGTTTGAATAATATCACTAAATGATCTGAGGCTGAAGGATTTGGGAAACATTTCTCAGACAT	1860	
1801	Db	AACTGTTTGAATAATATCACTAAATGATCTGAGGCTGAAGGATTTGGGAAACATTTCTCAGACAT	1860	
1861	Qy	TGAAAAATATCACTTTACTCCAAGTGCAGAAAGTCTCGCTGACTGGGAAATGACCA	1920	
1861	Db	TGAAAAATATCACTTTACTCCAAGTGCAGAAAGTCTCGCTGACTGGGAAATGACCA	1920	
1921	Qy	ACGATCTGGGCTCCAACATGACCATCGGGCTCTGAAACAGCAGAGGTGAATTCACAGGCA	1980	
1921	Db	ACGATCTGGGCTCCAACATGACCATCGGGCTCTGAAACAGCAGAGGTGAATTCACAGGCA	1980	
1981	Qy	CCTACATCAAGCGGTAAACAGCCACATCAATGAGATCAAAAGAGTCCACTGCATGGGA	2040	
1981	Db	CCTACATCAAGCGGTAAACAGCCACATCAATGAGATCAAAAGAGTCCACTGCATGGGA	2040	
2041	Qy	CACAAAACCATCAACAGAGAGCCAGCCACCTTTGGCTTCACCGTCAATTTGGAAGT	2100	
2041	Db	CACAAAACCATCAACAGAGAGCCAGCCACCTTTGGCTTCACCGTCAATTTGGAAGT	2100	
2101	Qy	TTTCAGAGTCCACCATGTCTTCAAGGCGCAGTGTCTTATAGACAGGAAATGGGAAGGAGG	2160	
2101	Db	TTTCAGAGTCCACCATGTCTTCAAGGCGCAGTGTCTTATAGACAGGAAATGGGAAGGAGG	2160	
2161	Qy	TCCTGAAGACCATGTGGCTGCTCGGTTCAAGTGTTAATGACATTTGTFGTATGACTGGAAAG	2220	
2161	Db	TCCTGAAGACCATGTGGCTGCTCGGTTCAAGTGTTAATGACATTTGTFGTATGACTGGAAAG	2220	
2221	Qy	CTACCAGGTCGGCATCAATCTTCACTCGCTCGGCGCACACAGAGGAGTGTAGTGATG	2280	
2221	Db	CTACCAGGTCGGCATCAATCTTCACTCGCTCGGCGCACACAGAGGAGTGTAGTGATG	2280	
2281	Qy	ACCAAGTCTCTCTGGACTCCAGGTGAAAGAGAGATAGAGGCCCTCTCTCGACAAATATGG	2340	
2281	Db	ACCAAGTCTCTCTGGACTCCAGGTGAAAGAGAGATAGAGGCCCTCTCTCGACAAATATGG	2340	
2341	Qy	TATACCAAGGCTTCCAGGTTAAATAGTACTCCAGGCTTTTAAAGGTTATCGGGGGATCT	2400	

Db 2341 TATACCAAGCTTTCCAGGCTCTAATAGGTACTCCAGGTCTTAAAGGTGATCGGGGGATCT 2400
Qy 2401 CTGGTTTACCTGGAGTTTCAGAGATTCACAGGACCAATGGGGAAGACCGGGAGCCAGGAC 2460
Db 2401 CTGGTTTACCTGGAGTTTCAGAGATTCACAGGACCAATGGGGAAGACCGGGAGCCAGGAC 2460
Qy 2461 TTAATGGCAAAAGCCAGAGGAGGAGAAAGAGGAGTGGAGCATGCAAGACAAATCTA 2520
Db 2461 TTAATGGCAAAAGCCAGAGGAGGAGAAAGAGGAGTGGAGCATGCAAGACAAATCTA 2520
Qy 2521 ATACAGTCCGACTCGTGGGTGGCAGCGCCCTCAGCAAGGACAGAGTGGAGATTTTTCAG 2580
Db 2521 ATACAGTCCGACTCGTGGGTGGCAGCGCCCTCAGCAAGGACAGAGTGGAGATTTTTCAG 2580
Qy 2581 AAGCCAGTGGGATACGTTGTGTGACAGACCGCTGGGAACTCGCTGGAGGACTGTCGTCT 2640
Db 2581 AAGCCAGTGGGATACGTTGTGTGACAGACCGCTGGGAACTCGCTGGAGGACTGTCGTCT 2640
Qy 2641 GCAGGAGCTTGGGATACAAAGGTGTTCAAGTGTGCATAAGCGAGCTTATTTTGGAAAAG 2700
Db 2641 GCAGGAGCTTGGGATACAAAGGTGTTCAAGTGTGCATAAGCGAGCTTATTTTGGAAAAG 2700
Qy 2701 GTACGGGTCCAAATATGGCTGAATGAAGTATTTTTCGGGAAGAGATCATCCATTGAAG 2760
Db 2701 GTACGGGTCCAAATATGGCTGAATGAAGTATTTTTCGGGAAGAGATCATCCATTGAAG 2760
Qy 2761 AGTCAGAAATTAGACAGTGGGTGTGAGAGCCTGTTCCAGCAGCAAGAGTGTGGGGTTC 2820
Db 2761 AGTCAGAAATTAGACAGTGGGTGTGAGAGCCTGTTCCAGCAGCAAGAGTGTGGGGTTC 2820
Qy 2821 ACTTTGCACTTACATAAGTGCATCATATTTTCATTCACATTTTAACTGTTTAAAGTG 2880
Db 2821 ACTTTGCACTTACATAAGTGCATCATATTTTCATTCACATTTTAACTGTTTAAAGTG 2880
Qy 2881 ATTTTTCCTTTGCTTCACTAAATCAGCTTAAATTAATTTTAAAGAACTTAAAGATTTT 2940
Db 2881 ATTTTTCCTTTGCTTCACTAAATCAGCTTAAATTAATTTTAAAGAACTTAAAGATTTT 2940
Qy 2941 ATCCAAGAAAGGAATATTTAAATCACTGGATAAATCAATATAAATAGCTTCAATTT 3000
Db 2941 ATCCAAGAAAGGAATATTTAAATCACTGGATAAATCAATATAAATAGCTTCAATTT 3000
Qy 3001 GCTTCAAAATACAGAACCATTTTCAACTCTCTAGGTTTTTAAAGTGGCTGCTGCGGAATTG 3060
Db 3001 GCTTCAAAATACAGAACCATTTTCAACTCTCTAGGTTTTTAAAGTGGCTGCTGCGGAATTG 3060
Qy 3061 ATCCCTCAGGATATAGTATTTTCTGCTTTTCATAGGAGGGGAAATGATGCTTATGC 3120
Db 3061 ATCCCTCAGGATATAGTATTTTCTGCTTTTGCATAGGAGGGGAAATGATGCTTATGC 3120
Qy 3121 AATACTCTTGTAGTCTTGCACATGTTAAACATGATGTTAGCAACATGCTTTACAGGAGA 3180
Db 3121 AATACTCTTGTAGTCTTGCACATGTTAAACATGATGTTAGCAACATGCTTTACAGGAGA 3180
Qy 3181 GAAAGACACCGTGCATCCGATTTGGTGGAGTGAAGTGAATGATGCTTATTTAGG 3240
Db 3181 GAAAGACACCGTGCATCCGATTTGGTGGAGTGAAGTGAATGATGCTTATTTAGG 3240
Qy 3241 AAGGCAACAGACGGGTCTGACATGATTTGGACGACCACTGAATTCGCGATTCAGAGAT 3300
Db 3241 AAGGCAACAGACGGGTCTGACATGATTTGGACGACCACTGAATTCGCGATTCAGAGAT 3300
Qy 3301 ATTGTATTTAAGTGTCTAGCTGCATACAGCAAAAGCCATTTGACCATTTCAACATTTGGT 3360
Db 3301 ATTGTATTTAAGTGTCTAGCTGCATACAGCAAAAGCCATTTGACCATTTCAACATTTGGT 3360
Qy 3361 GTGCACCTTCAAGCTTCACTGCTGCGCAAGCACTCAGGGGCTGCTAAAGGAAG 3420
Db 3361 GTGCACCTTCAAGCTTCACTGCTGCGCAAGCACTCAGGGGCTGCTAAAGGAAG 3420
Qy 3421 CGGAACAGTAGAAGCCAGTCCGCAAGAAACGGTGTGACCCCGGATGAATGTGAGCTAC 3480
Db 3421 CGGAACAGTAGAAGCCAGTCCGCAAGAAACGGTGTGACCCCGGATGAATGTGAGCTAC 3480

Qy 3481 TGGGCTATCTCGACAAGGGAAAAACGCAAGCCAAAGAGAAAGAGTGTAGCTTGCAGTGG 3540
Db 3481 TGGGCTATCTCGACAAGGGAAAAACGCAAGCCAAAGAGAAAGAGTGTAGCTTGCAGTGG 3540
Qy 3541 CTTACATGGCCATAGCTAGACTGGCGGTTTTTATGGACAGCAAGCAACCGGAATTCGCA 3600
Db 3541 CTTACATGGCCATAGCTAGACTGGCGGTTTTTATGGACAGCAAGCAACCGGAATTCGCA 3600
Qy 3601 GCTGGGCGCCCTCTGGTAAGGTTGGGAAGCCCTGCAAAAGTAAACTGATGGCTTTCTTG 3660
Db 3601 GCTGGGCGCCCTCTGGTAAGGTTGGGAAGCCCTGCAAAAGTAAACTGATGGCTTTCTTG 3660
Qy 3661 CCGCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCG 3720
Db 3661 CCGCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCG 3720
Qy 3721 TTTCCGATGATTTGAACAAGATGGAATGACACGAGTTCTCCGGCCGCTTGGGTGGAGAG 3780
Db 3721 TTTCCGATGATTTGAACAAGATGGAATGACACGAGTTCTCCGGCCGCTTGGGTGGAGAG 3780
Qy 3781 CTATTCGGCTATGATGGGCACAAACAGACAACTCGCTGCTCTGATGCCCGCTGTTCGG 3840
Db 3781 CTATTCGGCTATGATGGGCACAAACAGACAACTCGCTGCTCTGATGCCCGCTGTTCGG 3840
Qy 3841 CTGTACGCGCAGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTCCCTGAAT 3900
Db 3841 CTGTACGCGCAGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTCCCTGAAT 3900
Qy 3901 GAACTGACGACGAGGACGCGCGCTATCGTGGCTGGCCACGACGCGGCGTTCTTGGCGCA 3960
Db 3901 GAACTGACGACGAGGACGCGCGCTATCGTGGCTGGCCACGACGCGGCGTTCTTGGCGCA 3960
Qy 3961 GCTGTCTCGACGTTGTCACTGAAGCGGAAGGAGTGGCTGCTATTGGGCGAAGTGGCG 4020
Db 3961 GCTGTCTCGACGTTGTCACTGAAGCGGAAGGAGTGGCTGCTATTGGGCGAAGTGGCG 4020
Qy 4021 GGGCAGGATCTCTGTCTCATCTCACCTTCTCTCCGCAAGAAAGTATCCATCATGGCTGAT 4080
Db 4021 GGGCAGGATCTCTGTCTCATCTCACCTTCTCTCCGCAAGAAAGTATCCATCATGGCTGAT 4080
Qy 4081 GCAATGGCGGCTGCTATCGCTTGTATCGGCTACCTGCGCCATTCGACCAACCAAGCGAAA 4140
Db 4081 GCAATGGCGGCTGCTATCGCTTGTATCGGCTACCTGCGCCATTCGACCAACCAAGCGAAA 4140
Qy 4141 CATCGCATCGAGCAGCAGCTACTCGGATGGAAGCCGCTCTTGTGATCAGATCATCTG 4200
Db 4141 CATCGCATCGAGCAGCAGCTACTCGGATGGAAGCCGCTCTTGTGATCAGATCATCTG 4200
Qy 4201 GACGAAGACATCAGGGGCTCGCGCCAGCCGAACTGTTTCGCGAGGCTCAAGGCGCGCATG 4260
Db 4201 GACGAAGACATCAGGGGCTCGCGCCAGCCGAACTGTTTCGCGAGGCTCAAGGCGCGCATG 4260
Qy 4261 CCCGACGGGAGGATCTGCTGCTGACCCATGGCGATGCTGCTGCGCGCAATCATGCTG 4320
Db 4261 CCCGACGGGAGGATCTGCTGCTGACCCATGGCGATGCTGCTGCGCGCAATCATGCTG 4320
Qy 4321 GAAATGGCGCTTTCTGGAATTCATCGACTGTGGCGGCTGGGTGTTGGCGGACCGCTAT 4380
Db 4321 GAAATGGCGCTTTCTGGAATTCATCGACTGTGGCGGCTGGGTGTTGGCGGACCGCTAT 4380
Qy 4381 CAGGACATAGCTTGGCTTACCCGCTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGAC 4440
Db 4381 CAGGACATAGCTTGGCTTACCCGCTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGAC 4440
Qy 4441 CGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTCGCAAGCGATCGCTTCTATCGC 4500
Db 4441 CGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTCGCAAGCGATCGCTTCTATCGC 4500
Qy 4501 CTTCTTCGAGGTTCTTCTGAGCGGAGTCTGCGGTTTCGATAAAATAAAAGATTTATTT 4560
Db 4501 CTTCTTCGAGGTTCTTCTGAGCGGAGTCTGCGGTTTCGATAAAATAAAAGATTTATTT 4560

Qy	4561	AGTCTCAGAAAAAGGGGGGAATGAAGA	CCCAACCTGTAGTGTGGCAGACTGACTTAA	4620
Db	4561	AGTCTCAGAAAAAGGGGGGAATGAAGA	CCCAACCTGTAGTGTGGCAGACTGACTTAA	4620
Qy	4621	GTAACGCCATTTTCAAGGCATGCAAAAA	TACATAACTGAGAATAGAGAAGTTCCAGATCA	4680
Db	4621	GTAACGCCATTTTCAAGGCATGCAAAAA	TACATAACTGAGAATAGAGAAGTTCCAGATCA	4680
Qy	4681	AGGTCAAGAAACAGATGGAACAGCTGAA	TATATGGGCCAAACAGGATATCTGTGGTAAGCAGT	4740
Db	4681	AGGTCAAGAAACAGATGGAACAGCTGAA	TATATGGGCCAAACAGGATATCTGTGGTAAGCAGT	4740
Qy	4741	TCCTGCCCCGGCTCAGGGCCNAAGAAC	AGATGGAACAGCTGNAATATGGGCCAAACAGGATA	4800
Db	4741	TCCTGCCCCGGCTCAGGGCCNAAGAAC	AGATGGAACAGCTGNAATATGGGCCAAACAGGATA	4800
Qy	4801	TCGTGTGTAAGCAGTTCTCTGCCCGGCT	CAGGGCCAAAGAACAGATGCTGCCAGATCGGG	4860
Db	4801	TCGTGTGTAAGCAGTTCTCTGCCCGGCT	CAGGGCCAAAGAACAGATGCTGCCAGATCGGG	4860
Qy	4861	TCGAGCCCTCAGCAGTGTTCATAGAGAA	CGATCAGATGTTTCAGGGTGCCCCCAAGGACCTG	4920
Db	4861	TCCAGCCCTCAGCAGTGTTCATAGAGAA	CGATCAGATGTTTCAGGGTGCCCCCAAGGACCTG	4920
Qy	4921	AAATGACCTGTGCTTATTTTGAACCTAA	CAACTCAGTTTCGCTTCTCGCTTCTGTTCGCGC	4980
Db	4921	AAATGACCTGTGCTTATTTTGAACCTAA	CAACTCAGTTTCGCTTCTCGCTTCTGTTCGCGC	4980
Qy	4981	GCTTCTGCTCCCGAGCTCAATAAAGAGCC	CACAACCCCTCACTCGGGGGCGCCAGTCCCT	5040
Db	4981	GCTTCTGCTCCCGAGCTCAATAAAGAGCC	CACAACCCCTCACTCGGGGGCGCCAGTCCCT	5040
Qy	5041	CCGATTGACTGAGTGC	CGCGGTACCGTGATCCCAATAAACCCCTCTTGCAGTTGCATCC	5100
Db	5041	CCGATTGACTGAGTGC	CGCGGTACCGTGATCCCAATAAACCCCTCTTGCAGTTGCATCC	5100
Qy	5101	GACTTGTGGTCTCGCTGTTCTTGGGAGGG	TCTCCTCTGAGTGAATGACTACCCCGTCAGC	5160
Db	5101	GACTTGTGGTCTCGCTGTTCTTGGGAGGG	TCTCCTCTGAGTGAATGACTACCCCGTCAGC	5160
Qy	5161	GGGGGTCTTTCATTTGG	5177	
Db	5161	GGGGGTCTTTCATTTGG	5177	

RESULT 2

```

RES011 2
US-09-897-006--9
; Sequence 9, Application US/09897006
; Patent No. US20020106729A1
; GENERAL INFORMATION:
; APPLICANT: Black, Gregory
; TITLE OF INVENTION: Expression Vectors
; FILE REFERENCE: GALA-06415
; CURRENT APPLICATION NUMBER: US/09/897,006
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,851
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 5130
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-897-006--9

```

```

Query Match      30.7%; Score 1588.6; DB 3; Length 5130;
Best Local Similarity 66.1%; Pred. No. 0;
Matches 3007; Conservative 0; Mismatches 754; Indels 786; Gaps 18;

Qy      1 TTTGAAAGACCCCAACCGTAGTGGCAAGCTTAGTTAAGTAACGCCACTTTGCAAGGCAT 60

```

Db	1	TTTGAAGAACCCCAACCCCGTAGTGGCGAAGCTAGCTAAGTAAAGCCCACTTTGCAAGGCAT	60
Qy	61	GGAAAAATACATAAATCTGAGAAATAGAAAAGTTTCAGATCAAGGTCAAGAAACAAAGAAACAGC	120
Db	61	GGAAAAATACATAAATCTGAGAAATAGAAAAGTTTCAGATCAAGGTCAAGAAACAAAGAAACAGC	120
Qy	121	TGAATACCAACAGGATATCTGTGGTAAGCGGTTCCTGCCCCGGCTCAGGGCCAAAGAAC	180
Db	121	TGAATACCAACAGGATATCTGTGGTAAGCGGTTCCTGCCCCGGCTCAGGGCCAAAGAAC	180
Qy	181	GATGAGACGCTGATGATGGGCCCAACAGGATATCTGTGGTAAGCAGTTCCCTGCCCCCG	240
Db	181	GATGAGACGCTGATGATGGGCCCAACAGGATATCTGTGGTAAGCAGTTCCCTGCCCCCG	240
Qy	241	CTCGGGCCCAAGAACAGATGGTCCCAAGATGCGGTCCAGCCCTCAGCAGTTCTTAGTGAA	300
Db	241	CTCGGGCCCAAGAACAGATGGTCCCAAGATGCGGTCCAGCCCTCAGCAGTTCTTAGTGAA	300
Qy	301	TCATCAGATGTTTCCAGGGTGCCCAAGGACCTGGAATAAGACCCCTGTACTTATTGGAAC	360
Db	301	TCATCAGATGTTTCCAGGGTGCCCAAGGACCTGGAATAAGACCCCTGTACTTATTGGAAC	360
Qy	361	TAACCAATCAGTTGCGTCTTCGCTTCTGTTCGCGCGCTTCCGCTCTCCGAGCTCAATAA	420
Db	361	TAACCAATCAGTTGCGTCTTCGCTTCTGTTCGCGCGCTTCCGCTCTCCGAGCTCAATAA	420
Qy	421	AGACCCCAACACCCCTCACTCGCGCGGCCAGTCTTCCGATAGACTGCGTCGCCCGCGGTAC	480
Db	421	AGACCCCAACACCCCTCACTCGCGCGGCCAGTCTTCCGATAGACTGCGTCGCCCGCGGTAC	480
Qy	481	CCGTATTCCCAATAAAGCCTCTTGCTGTTTGGCATCCGAATCGTGGTCTCGCTGTTCTCTTG	540
Db	481	CCGTATTCCCAATAAAGCCTCTTGCTGTTTGGCATCCGAATCGTGGTCTCGCTGTTCTCTTG	540
Qy	541	GGAGGTCCTCTGAGTGATGACTACCAAGACGGGGTCTTTCAATTTGGGGGCTCGT	600
Db	541	GGAGGTCCTCTGAGTGATGACTACCAAGACGGGGTCTTTCAATTTGGGGGCTCGT	600
Qy	601	CCGGATTTTGGAGACCCCTCCCAAGGACCAACCGACCCACACCGGAGGTAGCTGGCC	660
Db	601	CCGGATTTTGGAGACCCCTCCCAAGGACCAACCGACCCACACCGGAGGTAGCTGGCC	660
Qy	661	AGCAACTTATCTGTGTCGTCCGATTTGCTAGTGTCTATGTTTGTATGTTATGCGCCTGCG	720
Db	661	AGCAACTTATCTGTGTCGTCCGATTTGCTAGTGTCTATGTTTGTATGTTATGCGCCTGCG	720
Qy	721	TCGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCCGTTGGTGGAACTGACAGTT	780
Db	721	TCGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCCGTTGGTGGAACTGACAGTT	780
Qy	781	CTGAACACCCGGCGCAACCTCGGAGAGCTCCAGGGACTTTGGGGCGGTTTTGTGG	840
Db	781	CTGAACACCCGGCGCAACCTCGGAGAGCTCCAGGGACTTTGGGGCGGTTTTGTGG	840
Qy	841	CCGACCTGAGGAAGGGAGTCGATGTGGATCCGACCCCGTCAAGATATGTGGTCTTGTT	900
Db	841	CCGACCTGAGGAAGGGAGTCGATGTGGATCCGACCCCGTCAAGATATGTGGTCTTGTT	900
Qy	901	AGGAGACGAAACCTTAAACAGTTTCCGCTCCGCTCTGAATTTTGGTTTCGGTTTGGAA	960
Db	901	AGGAGACGAAACCTTAAACAGTTTCCGCTCCGCTCTGAATTTTGGTTTCGGTTTGGAA	960
Qy	961	CCGAGCCGCGCTTGTCTGCTGAGCCAAAGCTTTGGGCTGAGCTCGACTCTAGAGGA	1020
Db	961	CCGAGCCGCGCTTGTCTGCTGAGCCAAAGCTTTGGGCTGAGCTCGACTCTAGAGGA	1020
Qy	1021	TCATTTCCGACAGGATAAATCGGTGCTTCCGCTTTTAGGACATATGAAGTATGGCACAGT	1080
Db	1021	CC-----CTGCACCGAACATGGAGAACACAAACATCAGGATTCCT	1059
Qy	1081	GGGATGATTTCTGTATCAGCAAGAGGACACTGACAGCTGTGACAGAGTCTGTGAAGTTCTG	1140
Db	1060	AGACCCCTGCTCGTGTTCAGCGGGGGTTTTTCTTTGTTGACAAGAATCTCTCAATATCC	1119

1141 ATGCTCGCTCAGTGACAGCTTTGCTTCTCTCCCAATCCTAAATAATGGGCCCAACTCTTCAAG 1200
1120 ACAGAGTCTAGAGTCTGCTGGTGGAGTCTCTCAATTTCTAGGGGAGCACCCAGTGTCC 1179
1201 AGAGGATGAAGTCTTATAAACTGACATGATCACCTTTTATCTCATTTGTTGTTGATGTC 1260
1180 TGGCCAAAATTCGGAGTCCCACTTCCCAATCACTACCAACCTCTTGTCTCTCAATTTG 1239
1261 TCGTGCCCATCAITTTGGCATAGTGGCAGCTCAGCTCTGAAATGGGAAACGAAAGATTGCA 1320
1240 TCCTGGCTATC-----GCTGGATGTCTCTGGGCGTTTATC 1276
1321 CGGTTGGCTCAGTTAATGCGAGATATCTCAGAGTCCGGAAGGCAAGAAATGGCAGTG 1380
1277 ATATTCTCTTCTCCTCTGCTATGCTCTATCTCTTGTGTTGTTCTTCTGGACTACCAA 1336
1381 AAGATGAATCAGATTTTCAGAGAGCTGTGAGGACGATGAGCAATGAGCAATGAGGAGCA 1440
1337 GGTATGTTGGCCGTTTGTCTCTACTTCCAGGAA---CATCAACTACAGCACGGGACCA 1393
1441 TCCAGTATCTTTCCAGATAATGAAGCCAACTCTCTAGATGCTAAGAAATTTCCAAAATTCA 1500
1394 TGGAGACCTGCGAGATTCCTGCTCAAGGAACCTCTATGTTTCCCTCTT-----GTTGCT 1448
1501 GCATAACAACTGATCAAGATTTAATGATGTTCTTTTCCAGCTAAATTTCTTACTTTCTCT 1560
1449 GTACAAAACCTTCGAGCGGAACTGCATTTGATTTCCCATCCATCTCTGGGCTTTG 1508
1561 CCATCCAGGAACATGAGNATATATAGGGGATATCTCCAGTCAATAGTAGTCTGAACA 1620
1509 CAGATTCCTATGGAGTGGCCCTCAGTCCGTTTCTCTCTG----- 1549
1621 CCACAGTACTTGATTTGAGTTTCAAGTTCAGTATTTGAAACACTGAATGGGAGTCCAGAGAA 1680
1550 -CTCAGTTTACTAGTGCCATTTGTTTCACTGTTTC---GTAGGGCTTTCCCACTGTTTG 1605
1681 CATTAAACAAAGAGGAGATGCGTAAATTAGAGGAGCGTATATACAATGATCATCAGCAG 1740
1606 GCTTTTCAATATG-GATGATGTTGATTTGGGGCCCAAGTCTGTACAACTCTTGATC 1664
1741 AAATTAAGTCTAGATGAAACAAAGTATATTTGGAACAGAAATTAAGGGGAAATGA 1800
1665 CTTTTTACCTCTATACCAATTTTCTTTTGTCTTGGGTATACATTTAAACCCATAA 1724
1801 AACTGTTGAATAATCACTAATGATCTGAGGCTGAGGATTTGGAAACATTTCTCAGACAT 1860
1725 AAC-----CAAGCT 1734
1861 TGAATAATATCACTTTTACTCCAAGGTGCGAGAAAGTGTCTGCTGACTGGGAAATGGACCA 1920
1735 TGGGGCTACTCCCTTAACCTTCACTGGGATATGTAATTTGGATTTGGG----- 1781
1921 ACGATCTGGGCTCCAAATGACATGACATCGGGGCTGTGAACAGCAGAGGTTGAATTCAGGCA 1980
1782 ----- 1781
1981 CCTATATCACAGCGTAAACGCCACATCAATGAGATCAAGAGTCAACCTGATGGGA 2040
1782 -----TACTTTTACCCAGAGAACATATGTAATAAATCAAGCAATGTTTTCG- 1829
2041 CACAAAACCAATCAACAAAGAGGAGCCAGCCACCTTTGGCTTCCACCGTCAATTTGGAAGT 2100
1830 -----AAAACCTGCTGTAAATAGACCTATTGATTTGGAAGTATGTCAGAGACTTTGGGT 1884
2101 TTTGAGTCCACATGCTCTTCAAGGCGCAGTGTCTCATAGACAGGAATGGGAAGG 2160
1885 CTTTGGGCTTTGTGCGCCCTTTTACAAATGTGGCTATCTGCTTAATGCCCTTTATAT 1944
2161 TCCTGAGACCACTGTGGCTGTGCGGTCAAGTGTAAATGACATTTGGTGATGACTGGAAG 2220
1945 GCATGTATACAACTCT----- 1959

2221 CTACAGGGTCGGCATCAACATCTTCACTCCCTGCGCACACACAGAGTGAAGTG 2280
1960 -----AAGCAGGCTTTTCACTTTCTCGCAACTTACAGGCTTTCTGTGTA 2005
2281 ACCAAGTCTCTCTGGACTCCAGGTGAAAAGGAGATAGAGGCCCTCTCTGGACAAAATGG 2340
2006 AACATATCTGAACCTTTACCCGTTGCCGGCA----- 2039
2341 TATACAGGCTTTTCCAGGCTCTAATAGGTACTTCCAGGTCTTTAAAGGTGATCGGGGGATCT 2400
2040 ----- 2039
2401 CTGGTTTACCTGGAGTTCGAGGATTCACAGACCAATGGGGAGACCGGGAGCCAGGAC 2460
2040 ----- 2039
2461 TTAATGACAAAAGGCCAGAGGAGGAGAAAAGGGAGTGGGAAGCATGCAAGACAAATCTA 2520
2040 -----ACGTCAG 2047
2521 ATACAGTCCGACTGGTGGTGGCGAGCGGCCCTCAAGAGGAGAGTGGAGATTTTTCACG 2580
2048 GTCTCTGCAAGTGTGCTGACGCAACCCCACTGGATGGGCTTGGCTATCGGCCATA 2107
2581 AAGCCAGTGGGGTACGCTGTGTGACGACCGCTGGGAACTGCGTGGAGGACTGGTCTGTCT 2640
2108 GCGCATGCGCGGACCTTTGTGGCTCTCTGCGC----- 2141
2641 GCAGGACTTTGGGATACAAAGGTGTTCAAGGTGTCATTAAGCGAGCTTATTTTCGAAAAG 2700
2142 -----ATCCATACTGCGGAACTCTCTAGCAGCTTGT----- 2171
2701 GTACGGGTCCAATATGCTGAATGAAGTATTTTGTTCGGGAAAGAGTCATCCATTGAAG 2760
2172 ----- 2171
2761 AGTGCAGAAATTAGACAGTGGGTGTGAGAGCTGTCGACGACGAGATGCTCGGGGTC 2820
2172 -----TTTGTCTCGCAGGCGGTCTGGAGCGAA 2197
2821 ACTTTGCACCTACATANTGATCATATTTTTCATTCACATTTTAAACTGTTATAAAGTG 2880
2198 ACTTATCGGC----- 2207
2881 ATTTTTTCTCTTCTTCACTAAATCAGCTTAATTAATATTTAAGAAACTAAGAAATTTT 2940
2208 ----- 2207
2941 ATCCACAGAAAAGGAATATTTAAATACTCTGGATAAACAATATAAATAGCTTCATATTT 3000
2208 -----ACCGACAACCTCT 2219
3001 GCTTCAATACCAGAACCATTTCACTTCTCTAGTGTTTTAAAGTGGCTCGTGGCGAATTTG 3060
2220 GTTGTCTCTCTCGGAATAACACTCTCTTCCATGCTGTAGGGTGTGCTGCCAATGG 2279
3061 ATCCCTCTCAGGATATAGTAGTTTTCGCTTTTGCATAGGAGGGGAAATGATGCTTATGC 3120
2280 ATCCCTCTCAGGATATAGTAGTTTTCGCTTTTGCATAGGAGGGGAAATGATGCTTATGC 2339
3121 AATACTCTTGTAGTCTTTGCAACATGTTAAACGATGAGTTAGCAACATGCTCTTACAGGAGA 3180
2340 AATACACTTGTAGTCTTTGCAACATGTTAAACGATGAGTTAGCAACATGCTCTTACAGGAGA 2399
3181 GAAAAGCACCGTGTGCTGCGATTTGGTGAAGTAAAGTGAATGATCGATCGTGGCTTATTAGG 3240
2400 GAAAAGCACCGTGTGCTGCGATTTGGTGAAGTAAAGTGAATGATCGATCGTGGCTTATTAGG 2459
3241 AAGCAACAGACGGCTGTGATGATTTGGACCAACCACTGAATTTCCGCACTTGAGAGAT 3300
2460 AAGCAACAGACGGCTGTGATGATTTGGACCAACCACTGAATTTCCGCACTTGAGAGAT 2519
3301 -ATTGTATTTAAGTGGCTGTGCTCGATACAGCAAAACGCCA--TTTGACCAATTCCACCACATT 3357

```
Db 2520 AATTGTTAATTAAAGTGCCTAGCTCGATACAGCAACGCCAATTTTGAACCATTCACCAAT 2579
Qy 3358 GGTGTGCACCT--CCAAGCTTCAGCTGCGCAGCACTCAGGGGCGCAAGGGCTGCTAAA 3415
Db 2580 GGTGTGCACCTTCCAAAGCTTCAGCTGCGCAGCACTCAGGGGCGCAAGGGCTGCTAAA 2639
Qy 3416 GGAAGCGGAACAAGTAGAAGCCAGTCCGAGAAACGGTGTGACCCCGGATGAATGTCA 3475
Db 2640 GGAAGCGGAACAAGTAGAAGCCAGTCCGAGAAACGGTGTGACCCCGGATGAATGTCA 2699
Qy 3476 GCTACTGGGCTATCTGGCAAGGGGAAACGCAAGCGCAAGAGAAAGCAGGTAGCTTGCA 3535
Db 2700 GCTACTGGGCTATCTGGCAAGGGGAAACGCAAGCGCAAGAGAAAGCAGGTAGCTTGCA 2759
Qy 3536 GTGGGCTTACATGGGATAGCTAGCTGGGCGGTTTATGACACAGCAAGCAACCGAAT 3595
Db 2760 GTGGGCTTACATGGGATAGCTAGCTGGGCGGTTTATGACACAGCAAGCAACCGAAT 2819
Qy 3596 TGCCAGCTGGGGCGGCTCTGTAAGTGTGGAAAGCCCTGCAAAAGTAACTGGATGCTT 3655
Db 2820 TGCCAGCTGGGGCGGCTCTGTAAGTGTGGAAAGCCCTGCAAAAGTAACTGGATGCTT 2879
Qy 3656 TCTTGCCGCCAAGGATCTGATGGGCGAGGGATCAAGATCTGATCAAGAGACAGGATGAG 3715
Db 2880 TCTTGCCGCCAAGGATCTGATGGGCGAGGGATCAAGATCTGATCAAGAGACAGGATGAG 2939
Qy 3716 GATGTTTCGATGATGAAACAAGATGATGACAGCGAGTTCTCCGGCGCTGGGTGG 3775
Db 2940 GATGTTTCGATGATGAAACAAGATGATGACAGCGAGTTCTCCGGCGCTGGGTGG 2999
Qy 3776 AGAGCTATTCCGCTATGACTGGGCACAAAGACATCGGCTGTCTGATCGCCCGGTGT 3835
Db 3000 AGAGCTATTCCGCTATGACTGGGCACAAAGACATCGGCTGTCTGATCGCCCGGTGT 3059
Qy 3836 TCCGCTGTACGCGAGGGGCGCGGTTCCTTTTGTCAAGACCGACTGTCGGTGCC 3895
Db 3060 TCCGCTGTACGCGAGGGGCGCGGTTCCTTTTGTCAAGACCGACTGTCGGTGCC 3119
Qy 3896 TGAATGAATCGAGACAGGAGCGCGGCTATCGTGCTGCGCACGAGCGGCTTCCTT 3955
Db 3120 TGAATGAATCGAGACAGGAGCGCGGCTATCGTGCTGCGCACGAGCGGCTTCCTT 3179
Qy 3956 GCGAGCTGTCTCGAGCTGTCTAAGCGGAGAGGAGTGTGCTGATTTGGGCGAG 4015
Db 3180 GCGAGCTGTCTCGAGCTGTCTAAGCGGAGAGGAGTGTGCTGATTTGGGCGAG 3239
Qy 4016 TGCGGGGCGAGGATCTCCTGTCACTCACTTGTCTTGTGCGAGAAAGTATCCATCATGG 4075
Db 3240 TGCGGGGCGAGGATCTCCTGTCACTCACTTGTCTTGTGCGAGAAAGTATCCATCATGG 3299
Qy 4076 CTGATGCAATCGGGGCGTGCATAGCTTGTATCGGCTACCTGCGCATTCGACCAACAG 4135
Db 3300 CTGATGCAATCGGGGCGTGCATAGCTTGTATCGGCTACCTGCGCATTCGACCAACAG 3359
Qy 4136 CGAAACATCGATCGAGCGACAGTACTCGGATGGAAGCGGCTTGTGCTCAGGATG 4195
Db 3360 CGAAACATCGATCGAGCGACAGTACTCGGATGGAAGCGGCTTGTGCTCAGGATG 3419
Qy 4196 ATCTGGACGAAGAGATCAGGGGCTCGCGCAGCGCAACTGTTCCGAGGCTCAAGGCGC 4255
Db 3420 ATCTGGACGAAGAGATCAGGGGCTCGCGCAGCGCAACTGTTCCGAGGCTCAAGGCGC 3479
Qy 4256 GCATGCCCGAGCGGAGATCTGCTGTGACCCATGCGGATGCTGTTGCCGAATATCA 4315
Db 3480 GCATGCCCGAGCGGAGATCTGCTGTGACCCATGCGGATGCTGTTGCCGAATATCA 3539
Qy 4316 TGGTGGAATAGCGGCTTTCTGATTCATCGACTGTGGCGGCTGGGTGCGGAC 4375
Db 3540 TGGTGGAATAGCGGCTTTCTGATTCATCGACTGTGGCGGCTGGGTGCGGAC 3599
Qy 4376 GCTATCAGACATAGCTGTGGCTACCCGCTGATATTGCTGAAGAGCTTGGCGGCAATGGG 4435
```

```
Db 3600 GCTATCAGGACATAGCGTTGGCTACCGGTGATATTGCTGAAGAGCTTTGGCGCGAATGGG 3659
Qy 4436 CTGACCGCTTCCCTCGTCTTTTACGGTATCGCGCTCCCGATTTCGACGCGATCGCTTCT 4495
Db 3660 CTGACCGCTTCCCTCGTCTTTTACGGTATCGCGCTCCCGATTTCGACGCGATCGCTTCT 3719
Qy 4496 ATGCGCTTCTTGACGAGTCTTCTTGAGCGGAGCTCTCGGGGTTTCGATA 4542
Db 3720 ATGCGCTTCTTGACGAGTCTTCTTGAGCGGAGCTCTCGGGGTTTCGATA 3766

RESULT 3
US-09-897-511A-9
; Sequence 9, Application US/09897511A
; Publication No. US20030092882A1
; GENERAL INFORMATION:
; APPLICANT: Bremel, Robert
; APPLICANT: Miller, Linda
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors
; FILE REFERENCE: GALA-06416
; CURRENT APPLICATION NUMBER: US/09/897,511A
; CURRENT FILING DATE: 2001-06-29
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 5130
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-897-511A-9

Query Match 30.7%; Score 1588.6; DB 3; Length 5130;
Best Local Similarity 66.1%; Pred. No. 0;
Matches 3007; Conservative 0; Mismatches 754; Indels 786; Gaps 18;

Qy 1 TTTGAAAGACCCACCGTAGGTGGCAAGCTTAGTAAAGCCACTTTGCAAGGCAT 60
Db 1 TTTGAAAGACCCACCGTAGGTGGCAAGCTTAGTAAAGCCACTTTGCAAGGCAT 60
Qy 61 GGAATAATACATACTAGTATAGAAAGTTTCAAGTCAAGGTCAAGAAAGAAACAGC 120
Db 61 GGAATAATACATACTAGTATAGAAAGTTTCAAGTCAAGGTCAAGAAAGAAACAGC 120
Qy 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTTCTGCCCCGGCTCAGGGCCAAAGACA 180
Db 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTTCTGCCCCGGCTCAGGGCCAAAGACA 180
Qy 181 GATGAGACAGCTGATGATGGGCAACAGGATATCTGTGTAAGCGGTTTCTGCCCCGG 240
Db 181 GATGAGACAGCTGATGATGGGCAACAGGATATCTGTGTAAGCGGTTTCTGCCCCGG 240
Qy 241 CTGGGGCCCAAGAACAGATGTCCTCCAGATGCGGTCAGCCCTCAGCAGTTTCTAGTGA 300
Db 241 CTGGGGCCCAAGAACAGATGTCCTCCAGATGCGGTCAGCCCTCAGCAGTTTCTAGTGA 300
Qy 301 TCATCAGATGTTTCCAGGGTGCCTCCAGGACCTGAAATGACCTGTACCTTATTTGAAC 360
Db 301 TCATCAGATGTTTCCAGGGTGCCTCCAGGACCTGAAATGACCTGTACCTTATTTGAAC 360
Qy 361 TAACCAATCAGTTCGCTTCTGCTTCTGTCGCGCTCTCGGCTCTCGAGCTCAATAA 420
Db 361 TAACCAATCAGTTCGCTTCTGCTTCTGTCGCGCTCTCGGCTCTCGAGCTCAATAA 420
Qy 421 AGAGCCCAACACCTCTCAGTCGCGCGCCAGTCTTCGATAGACTGCGTCCCGGGTAC 480
Db 421 AGAGCCCAACACCTCTCAGTCGCGCGCCAGTCTTCGATAGACTGCGTCCCGGGTAC 480
Qy 481 CCGTATTTCCCAATAAAGCCTCTTCTGTTTTCGATCCGAATCGTGTCTCGCTTCTTGT 540
```


QY 2701 GTACGGSTCCAAATATGCTGAATGAAGTATTTTGTTCGGGAAAGATCATCCATTGAAG 2760
Db 2172 ----- 2171
QY 2761 AGTCAGAAATTAGACAGTGGGGTGTGAGAGCTGTTCCACGACGAAAGATCTCGGGGTC 2820
Db 2172 -----TTTGCTCGCAGCGGCTGTGGAGCGAA 2197
QY 2821 ACTTTGCACCTACATAAATGATCATATTTTTCATTCACATTTTAAAACTGTTATAAAGTG 2880
Db 2198 ACTTATCGC----- 2207
QY 2881 ATTTTTCCTTTGCTTCACTAAATCAGCTTAATTAATAATTAAGAAACTAAGAAATTTT 2940
Db 2208 ----- 2207
QY 2941 ATCCACAGAAAGGAATATTTAAATAATCACTGGATAAATACATAATAAATAGCTTCATATTT 3000
Db 2208 -----ACCGACAACCTCT 2219
QY 3001 GCTTCANATACAGAACCATTTCAACTTCTCTAGTCTTTTAAAGTGGCTCGTGCCGAATTG 3060
Db 2220 GTTGTCTCTCTCGGAATATACACTCTCTTTCATGGCTGTGCTGAGGTGTGCTGCCAATGG 2279
QY 3061 ATCCCTCAGGATATAGTAGTTTCGCTTTTGCATAGGGAGGGGAAATGTAGTCTTATGC 3120
Db 2280 ATCCCTCAGGATATAGTAGTTTCGCTTTTGCATAGGGAGGGGAAATGTAGTCTTATGC 2339
QY 3121 AATACTCTGTAGTCTTCAACATGGTAAACGATGAGTTAGCAACATGCTTTTACAAGGAGA 3180
Db 2340 AATACTCTGTAGTCTTCAACATGGTAAACGATGAGTTAGCAACATGCTTTTACAAGGAGA 2399
QY 3181 GAAAGACACCGTGCATGCCATTTGGTGAAGTAAGTGGTACGATCGTGCTTTATTAGG 3240
Db 2400 GAAAGACACCGTGCATGCCATTTGGTGAAGTAAGTGGTACGATCGTGCTTTATTAGG 2459
QY 3241 AAGGCAACAGACGGGCTCTGACATGATTTGAGACGAAACCACTGAAATTCGCGATTCAGAGAT 3300
Db 2460 AAGGCAACAGACGGGCTCTGACATGATTTGAGACGAAACCACTGAAATTCGCGATTCAGAGAT 2519
QY 3301 -ATTGATTTAAGTGCCTAGCTCGATACAGCAAAACGCCA--TTTGACCAATTCACCAAT 3357
Db 2520 AATTGATTTAAGTGCCTAGCTCGATACAGCAAAACGCCAATTTTGAACCAATTCACCAAT 2579
QY 3358 GGTGTGCACCT--CCAGCTTCAGCTGCGCAAGCACTCAGGGGCGAAGGGCTCTAAA 3415
Db 2580 GGTGTGCACCTTCCAAAGCTTCAGCTGCGCAAGCACTCAGGGGCGAAGGGCTCTAAA 2639
QY 3416 GGAAGCGGAACACGTAGAAAGCCAGTCCGCAAGAACCGGTGCTGACCCCGGATGAATGTCA 3475
Db 2640 GGAAGCGGAACACGTAGAAAGCCAGTCCGCAAGAACCGGTGCTGACCCCGGATGAATGTCA 2699
QY 3476 GCTACTGGGCTATCTGGCAAGGGGAAACGCAAGCGCAAGAGAAAGCAGGTAGCTTGCA 3535
Db 2700 GCTACTGGGCTATCTGGCAAGGGGAAACGCAAGCGCAAGAGAAAGCAGGTAGCTTGCA 2759
QY 3536 GTGGGCTTACATGGGATAGCTAGCTGGGCGGTTTATGACACAGCAAGCAAGCGGAAT 3595
Db 2760 GTGGGCTTACATGGGATAGCTAGCTGGGCGGTTTATGACACAGCAAGCAAGCGGAAT 2819
QY 3596 TGCCAGCTGGGCGGCTCTGTAAGTTGGGAAGCCCTGCAAAAGTAACCTGATGGCTT 3655
Db 2820 TGCCAGCTGGGCGGCTCTGTAAGTTGGGAAGCCCTGCAAAAGTAACCTGATGGCTT 2879
QY 3656 TCTTGCCGCCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGGATGAG 3715
Db 2880 TCTTGCCGCCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGGATGAG 2939
QY 3716 GATCGTTTCGATGATGAAACAAGATGATTCACGAGGTTCTCCGGCCCGCTTGGGTGG 3775
Db 2940 GATCGTTTCGATGATGAAACAAGATGATTCACGAGGTTCTCCGGCCCGCTTGGGTGG 2999
QY 3776 AGAGGCTATTCGGCTATGACTGGGCGACAAACAGACAAATCGGCTGCTCTGATGCCCGCTGT 3835

Db 3000 AGAGGCTATTCGGCTATGACTGGGCACAAACAGACAAATCGGCTGCTCTGATGCCCGCTGT 3059
QY 3836 TCCGGCTGTACGCGCAGGGGCGCCGGTCTTTTGTTCAGAGCAGCTGTCCGGTGCCC 3895
Db 3060 TCCGGCTGTACGCGCAGGGGCGCCGGTCTTTTGTTCAGAGCAGCTGTCCGGTGCCC 3119
QY 3896 TGAATGAACCTGCAGACGAGCAGCGCGCTATCGTGGCTGGCCACACGCGGCGTTCCTT 3955
Db 3120 TGAATGAACCTGCAGACGAGCAGCGCGCTATCGTGGCTGGCCACACGCGGCGTTCCTT 3179
QY 3956 GCGCAGCTGTCTGCAGCTTTGTCACTGAAGCGGAAAGGACTGTGCTGCTATTGGCGGAAG 4015
Db 3180 GCGCAGCTGTCTGCAGCTTTGTCACTGAAGCGGAAAGGACTGTGCTGCTATTGGCGGAAG 3239
QY 4016 TGCCGGGCGAGGATCTCTGCTCATCTCACCCTTGTCTCGGAGAAAGTATCCATCATGG 4075
Db 3240 TGCCGGGCGAGGATCTCTGCTCATCTCACCCTTGTCTCGGAGAAAGTATCCATCATGG 3299
QY 4076 CTGATGCAATGCGCGGCTGCATACGCTTGTATCGGCTTACTGCGCAATTCGACCACCAAG 4135
Db 3300 CTGATGCAATGCGCGGCTGCATACGCTTGTATCGGCTTACTGCGCAATTCGACCACCAAG 3359
QY 4136 CGAAACATCGCATCGAGCGAGCAGTACTCGGATGGAAGCGGCTTTGTGATCAGGATG 4195
Db 3360 CGAAACATCGCATCGAGCGAGCAGTACTCGGATGGAAGCGGCTTTGTGATCAGGATG 3419
QY 4196 ATCTGGAAGAAAGACATCAGGGGCTCGCGCAGCGCAACTGTTGCGCAGGCTCAAGGCGC 4255
Db 3420 ATCTGGAAGAAAGACATCAGGGGCTCGCGCAGCGCAACTGTTGCGCAGGCTCAAGGCGC 3479
QY 4256 GCATGCCCGACGCGAGGATCTCGTCTGAGCCCAATGCGGATGCTGCTGCGCAATATCA 4315
Db 3480 GCATGCCCGACGCGAGGATCTCGTCTGAGCCCAATGCGGATGCTGCTGCGCAATATCA 3539
QY 4316 TGGTGGAAATGCGCGCTTTTCTGGATTTCATCGACTGTGCGCGGCTGGGTGTGCGGAC 4375
Db 3540 TGGTGGAAATGCGCGCTTTTCTGGATTTCATCGACTGTGCGCGGCTGGGTGTGCGGAC 3599
QY 4376 GCTATCAGGACATAGCTTGGCTTACCGGTGATATTGCTGAAGAGCTTGGCGGCGCAATGG 4435
Db 3600 GCTATCAGGACATAGCTTGGCTTACCGGTGATATTGCTGAAGAGCTTGGCGGCGCAATGG 3659
QY 4436 CTGACCGCTTCTCGTCTTTTACGGTATCGCGCTCCCGATTCGACGCGCATCGCCTTCT 4495
Db 3660 CTGACCGCTTCTCGTCTTTTACGGTATCGCGCTCCCGATTCGACGCGCATCGCCTTCT 3719
QY 4496 ATCGCTTCTTGACGAGTCTTCTTGAGCGGAGCTCTGGGGTTCGATA 4542
Db 3720 ATCGCTTCTTGACGAGTCTTCTTGAGCGGAGCTCTGGGGTTCGATA 3766

RESULT 4

US-10-397-079-9
; Sequence 9, Application US/10397079
; Publication No. US20030224415A1
; GENERAL INFORMATION:
; APPLICANT: Bremel, Robert
; APPLICANT: Miller, Linda
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors
; FILE REFERENCE: GALA-06416
; CURRENT APPLICATION NUMBER: US/10397,079
; CURRENT FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: US/09/897,511A
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,925
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 5130
; TYPE: DNA

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-10-397-079-9

Query Match 30.7%; Score 1588.6; DB 6; Length 5130;
Best Local Similarity 66.1%; Pred No. 0;
Matches 3007; Conservative 0; Mismatches 754; Indels 786; Gaps 18;

Qy 1 TTTGAAGACCCACCCCTAGTGTGCAAGCTAGCTTAAAGTAAAGCAAGCTTTGCAAGGCAT 60
Db TTTGAAGACCCACCCCTAGTGTGCAAGCTAGCTTAAAGTAAAGCAAGCTTTGCAAGGCAT 60

Qy 61 GGAAAAATACATACTAGATAGATAAGAAAAGTTTCAAGTCAAGGTCAGGAACAAGAAACAGC 120
Db GGAAAAATACATACTAGATAGATAAGAAAAGTTTCAAGTCAAGGTCAGGAACAAGAAACAGC 120

Qy 121 TGAATACCAACAGGATATCTGTGTAGCGGTTCCTGCCCGGCTCAGGGCCAGAAACA 180
Db TGAATACCAACAGGATATCTGTGTAGCGGTTCCTGCCCGGCTCAGGGCCAGAAACA 180

Qy 181 GATGAGACAGCTGAGTGTGAGGCAAGCAAGATATCTGTGTAGCGGTTCCTGCCCGG 240
Db GATGAGACAGCTGAGTGTGAGGCAAGCAAGATATCTGTGTAGCGGTTCCTGCCCGG 240

Qy 241 CTGGGGCCAAAGACAGATGTTCCCAAGTCCAGCGGTCCAGCCCTCAGCAGTTTCTAGTGAA 300
Db CTGGGGCCAAAGACAGATGTTCCCAAGTCCAGCGGTCCAGCCCTCAGCAGTTTCTAGTGAA 300

Qy 301 TCATCAGATGTTCCAGGGTCCCAAGGACCTGAAATGACCTGTACCTTATTGAAAC 360
Db TCATCAGATGTTCCAGGGTCCCAAGGACCTGAAATGACCTGTACCTTATTGAAAC 360

Qy 361 TAAACCAATCAGTCTCGTCTCTGTTCTGCGCGCTTCGCGCTCCGAGCTCAATAAA 420
Db TAAACCAATCAGTCTCGTCTCTGTTCTGCGCGCTTCGCGCTCCGAGCTCAATAAA 420

Qy 421 AGAGCCCAACCCCTCACTCGGGGCCAGCTGTTCCGATAGACTGCTCCCGGGTAC 480
Db AGAGCCCAACCCCTCACTCGGGGCCAGCTGTTCCGATAGACTGCTCCCGGGTAC 480

Qy 481 CCGTATTCCTCAATAAGCTCTTGTGTTTGCATCCGAATCGTGTCTCGTGTCTCTTG 540
Db CCGTATTCCTCAATAAGCTCTTGTGTTTGCATCCGAATCGTGTCTCGTGTCTCTTG 540

Qy 541 GGAGGCTCTCTCAGTGTATGACTACCCAGCGGGGCTCTTCAATTTGGGGCTCGT 600
Db GGAGGCTCTCTCAGTGTATGACTACCCAGCGGGGCTCTTCAATTTGGGGCTCGT 600

Qy 601 CCGGATTTGGAGACCCCTGCCAGGGACCAAGCCACCAAGGAGGTAAGCTGGCC 660
Db CCGGATTTGGAGACCCCTGCCAGGGACCAAGCCACCAAGGAGGTAAGCTGGCC 660

Qy 661 AGCAATATCTGTCTGTCGATGTTGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 720
Db AGCAATATCTGTCTGTCGATGTTGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 720

Qy 721 TCTGTACTAGTTAGCTACTGATCTGTATCTGGCGGACCCGTGTGAACTGACGAGTT 780
Db TCTGTACTAGTTAGCTACTGATCTGTATCTGGCGGACCCGTGTGAACTGACGAGTT 780

Qy 781 CTGAACACCCCGCGCAACCCCTGGGAGACGTCCTCAGGGACCTTTGGGGCCGTTTGTGG 840
Db CTGAACACCCCGCGCAACCCCTGGGAGACGTCCTCAGGGACCTTTGGGGCCGTTTGTGG 840

Qy 841 CCCGACCTTGAGAAAGGAGTGTGATGTAATCCCGACCCCGTCAGGATATGTTGTTCTGGT 900
Db CCCGACCTTGAGAAAGGAGTGTGATGTAATCCCGACCCCGTCAGGATATGTTGTTCTGGT 900

Qy 901 AGGACACGAGACCTTAAACAGTTTCCCGCTCCGCTGTAATTTTGTCTTTCGTTTGAA 960
Db AGGACACGAGACCTTAAACAGTTTCCCGCTCCGCTGTAATTTTGTCTTTCGTTTGAA 960

Qy 961 CCGAAGCCGCGGCTTGTCTGCTGACGCCAAGCTTGGGCTGACGCTCTAGAGGA 1020
Db CCGAAGCCGCGGCTTGTCTGCTGACGCCAAGCTTGGGCTGACGCTCTAGAGGA 1020

Qy 1021 TCAATTCGGCACGAGTAAATCGGTCTCTTTAGGACATATGAAGTATGGCACAGT 1080
Db TCAATTCGGCACGAGTAAATCGGTCTCTTTAGGACATATGAAGTATGGCACAGT 1080

Qy 1081 GGGATGACTTTCTGATCAGCAAGGACACCTGACAGCTGTACAGAGTCTGTGAAGTTGG 1140
Db GGGATGACTTTCTGATCAGCAAGGAGGTTTTCTTGTGTGACAGAACTCTCACAATACC 1119

Qy 1141 ATGCTCGCTCAGTGCAGCTTTGCTCTCCCATCTCTAAATGGCCCAACTCTTCAAG 1200
Db ATGCTCGCTCAGTGCAGCTTTGCTCTCCCATCTCTCTCAATTTCTAGGGGAGACCCACGCTGCC 1179

Qy 1201 AGAGATGAACTCTTATAAAATGCACTGATGATCACCCTTTATCTATCTGTTGTAGTTC 1260
Db AGAGATGAACTCTTATAAAATGCACTGATGATCACCCTTTATCTATCTGTTGTAGTTC 1260

Qy 1261 TCGTGCCCATCAATTTGGCATAGTGGCAGCTCAGCTCTCTGAAATGGGAAACGAAATGCA 1320
Db TCGTGCCCATCAATTTGGCATAGTGGCAGCTCAGCTCTCTGAAATGGGAAACGAAATGCA 1320

Qy 1321 CCGTGTGGCTCAGTTAAATGACAGATATATCTCCAAGTCCGGAAGGCAAGAAATGGCAGTG 1380
Db CCGTGTGGCTCAGTTAAATGACAGATATATCTCCAAGTCCGGAAGGCAAGAAATGGCAGTG 1380

Qy 1381 AAGATGAAATGAGATTTTCAGAACTGTGAGGACGATGAGCAACATGGAAGACAGAA 1440
Db AAGATGAAATGAGATTTTCAGAACTGTGAGGACGATGAGCAACATGGAAGACAGAA 1440

Qy 1441 TCCAGTATCTTTTCCAGATAATGAAGCCATCTCTAGATGCTAAGAAATTTCCAAAATTTCA 1500
Db TCCAGTATCTTTTCCAGATAATGAAGCCATCTCTAGATGCTAAGAAATTTCCAAAATTTCA 1500

Qy 1501 GCATTAACAACTGATCAAGATTTAATGATGTTCTTTTCCAGTAAATTTCTTACTTCTCT 1560
Db GCATTAACAACTGATCAAGATTTAATGATGTTCTTTTCCAGTAAATTTCTTACTTCTCT 1560

Qy 1561 CCATCCAGGACATGAGATATCATAGGGATATCTCCAAGTCTCATCTAGTCTGAAACA 1620
Db CCATCCAGGACATGAGATATCATAGGGATATCTCCAAGTCTCATCTAGTCTGAAACA 1620

Qy 1621 CCAGTACTTGTATTTGAGTTCAGTATTTGAAACACTGAAATGGCAGAGTCCAAAGAAATG 1680
Db CCAGTACTTGTATTTGAGTTCAGTATTTGAAACACTGAAATGGCAGAGTCCAAAGAAATG 1680

Qy 1681 CATTTAAACAAAGAGAGATGCGTAAATTTAGAGGAGCGTATATACAAATGCATCAGCAG 1740
Db CATTTAAACAAAGAGAGATGCGTAAATTTAGAGGAGCGTATATACAAATGCATCAGCAG 1740

Qy 1741 AATTAAGTCTCTAGATGAAACAAAGTATATTTTGGAAACAGAAATAAAGGGGAAATGA 1800
Db AATTAAGTCTCTAGATGAAACAAAGTATATTTTGGAAACAGAAATAAAGGGGAAATGA 1800

Qy 1801 AACTGTTGAATAATCACTAATGATCTGAGGCTGAAGGATTTGGGAACTTCTCAGACAT 1860
Db AACTGTTGAATAATCACTAATGATCTGAGGCTGAAGGATTTGGGAACTTCTCAGACAT 1860

Qy 1861 TGAATAATATCACTTTTACTCCAAGTGCAGAAAGTGTCTGCTGATGGGAAATGGACCA 1920
Db TGAATAATATCACTTTTACTCCAAGTGCAGAAAGTGTCTGCTGATGGGAAATGGACCA 1920

Qy 1921 AGATCTGGGCTCCAACTGACCATCGGGCTGTGAACAGCAGAGGTTGAATTTCAAGGCA 1980
Db AGATCTGGGCTCCAACTGACCATCGGGCTGTGAACAGCAGAGGTTGAATTTCAAGGCA 1980

Qy 1981 CCTACATCACAGCGGTAAAGCCACATCAATGAGATCAAGAGTTCACCACTGCATGGGA 2040
Db CCTACATCACAGCGGTAAAGCCACATCAATGAGATCAAGAGTTCACCACTGCATGGGA 2040

Qy 2041 CACAAAACACCATCAACAGAGGAGGCCAGGCCACCTTTGGCTTCCACCGTCAATTTGGAA 2100
Db CACAAAACACCATCAACAGAGGAGGCCAGGCCACCTTTGGCTTCCACCGTCAATTTGGAA 2100

Db 1394 TCGAAGACCTGCACGATTCCTGCTCAAGGAACCTCTAATGTTTCCCTCTT-----GTTGCT 1448
Qy 1501 GGATAAACATGATCAAGAGTTAATGATGTTCTTTTCCAGGTAATTCCTTACTTTCCCT 1560
Db 1449 GTACAAAACCTTCGGACGGAACACTGCATCTGTATTTCCCATCCATCATCTCGGGCTTCG 1508
Qy 1561 CCATCCAGGACATGAGATATCATAGGGGATATCTCAAGTCATTAGTAGGTCGTGACA 1620
Db 1509 CAAGATTCCTATGGGAGTGGGCTCAGTCCGTTTCTCTCG----- 1549
Qy 1621 CCACAGTACTTGAATTTGAGTTTCAGATTGAAACACTGAATGGCAGAGTCCAAAGAGAATG 1680
Db 1550 -CTCAGTTTACTAGTGCCATTGTTTCAGTGGTTC---GTAGGGCTTCCCCCACTGTTG 1605
Qy 1681 CATTTAAACAACAAGAGGAGATGCGTAATATAGGAGCGGTATATACAAATGCATCAGCAG 1740
Db 1606 GCTTTTCAGTTATATG-GATGATGTGGTATTTGGGGCCAAAGTCTGTACAAACATCTTGAGTC 1664
Qy 1741 AAATTAAGTCTCTAGATGAAACAAAGTATATTTTGGACAGGAAATAAAGGGGAATGA 1800
Db 1665 CTTTTTACCTCTATTACCAATTTTCTTTGCTTTGGGTATACATTTAAACCCCTAATAA 1724
Qy 1801 AACTGTGAATAATATCACTAATGATCTGAGGCTGAAGGATTTGGGAACATTTCTCAGACAT 1860
Db 1725 AAC----- 1734
Qy 1861 TGAATAATATCATTTTACTTCCAGGTGCCAGAAAGTGTCTGCTGACTGGGAAATGGACCA 1920
Db 1735 TGGGCTACTCTCTTAACTTCATGGGATATGTAATTTGGATGTTGGG----- 1781
Qy 1921 ACGATCTGGGCTCAAACATGATGCCATCGGGGCTGTGACAGCAGAGGTGAATTCACAGGCA 1980
Db 1782 ----- 1781
Qy 1981 CCTACATCACAGCCGTAAACAGCCATCAATAGATCAAGAGTCAAGAGTCACTGCATGGGA 2040
Db 1782 -----TACTTTACCGCAAGACATATTTGCTACTTAAATTCAGCAATGTTTCG- 1829
Qy 2041 CACAAAACCATCAACAGAGGAGCCAGCCACCTTTGGTTTCCCGTCAATTTGGAAGT 2100
Db 1830 -----AAAATGCTGTAAATAGACTATTGATTTGAAAGTATGTACAGACTTTGTTGGT 1884
Qy 2101 TTTGAGTCCACACTGTTTTCAGGCGCCAGTGTCTCATAGACAGGAATGGGAGGAGG 2160
Db 1885 CTTTGGGCTTGTGCTCCCTTTTACAAATGTGGCTATCTGCTTAAATGCTTTATAT 1944
Qy 2161 TCCTGAAGACCATGTGGTGTCTGCTCGGTCAAGTGTAAATGACATTTGGTATGACTGGAAAG 2220
Db 1945 GCATGTATCAATCT----- 1959
Qy 2221 CTACAGGGTCGGCATCAACATCTTCTCTGCTCGGCACACAGAGGAGTGAAGTGTG 2280
Db 1960 -----AAGCAGGCTTTTCACTTTCTGCCCAACTTACAAAGCCCTTTCTGTGTA 2005
Qy 2281 ACCAAGTCTCTCTGAGTCCAGGTGAAAGAGGAGATAGAGCCCTCTCTGACAAATGG 2340
Db 2006 AACAAATATCTGAACCTTTTACCCCGTTGCCCGCA----- 2039
Qy 2341 TATACAGGCTTTCCAGGTCTAATAGTACTCCAGTCTTAAAGGTGATCGGGGGATCT 2400
Db 2040 ----- 2039
Qy 2401 CTGTTTACTGAGTTCAGGATTTCCAGGACCAATGGGGAAGACCGGGAAGCCAGGAC 2460
Db 2040 ----- 2039
Qy 2461 TTAATGGACAAAAGGCCAGAGGAGGAGAAAAGGAGTGGAGCATGCCAAGACATCTTA 2520
Db 2040 -----ACGTCAG 2047
Qy 2521 ATACAGTCCGACTGTGTGGTGGCAGCGCCCTCAGAAAGGACAGAGTGGAGATTTTTCAG 2580

Db 2048 GTCTCTGCCAAGTGTGTTGCTGACGCAACCCCACTGATGGGGCTTGCTATCGGCCATA 2107
Qy 2581 AAGCCAGTGGGTACGGTGTGTGACGACCGCTGGGAACCTGCGTGGAGGACTGCTGCT 2640
Db 2108 GCGCATGCGCGGACCTTTTGTGCTCTCTGCG----- 2141
Qy 2641 GCAGGAGCTTTGGGATACAAAGGTGTTCAAAAGTGTGCATAAGCGAGCTTATTTTGGAAAAG 2700
Db 2142 -----ATCCATCTCTGCGAATCTCTAGCAGCTGT----- 2171
Qy 2701 GTACGGTCCAAATATGCTGAATGAAGTATTTTGTGTTGGGAAAAGATCATTCATTTGAAG 2760
Db 2172 ----- 2171
Qy 2761 AGTCAGAAATTAGACAGTGGGGTGTGAGAGCCTGTTTCGACGACGAAGATGCTGGGGTC 2820
Db 2172 -----TTTGCTCGCAGGGGTCTGAGCGAA 2197
Qy 2821 ACTTTGCACCTACATAATGCATCATATTTTCATTCACATTTTTTAAACTGTTATAAAGTG 2880
Db 2198 ACTTATCGG----- 2207
Qy 2881 ATTTTTCCTTTGCTTCTCACTAAATCAGCTTAATAATATTTAAGAACTAAGAAATTTT 2940
Db 2208 ----- 2207
Qy 2941 ATCCACAGAAAGGAATATTTAAATACTGGAATAACATATAAATAGCTTCATATTT 3000
Db 2208 -----ACGCAACACTCT 2219
Qy 3001 GCTTCAAATACCAAGACCATTTTCAACTTCTTAGTGTGTTTAAAGTGGCTGCGCCGGAATG 3060
Db 2220 GTGTGCTCTCTCGGAAATACACCTCTTTCATGGCTGTGAGGTGTGCTGCCAAGCTGG 2279
Qy 3061 ATCCCTCAGGATATAGTAGTTTGCCTTTTCATAGGAGGGGGAATGTAGTCTTATGC 3120
Db 2280 ATCCCTCAGGATATAGTAGTTTGCCTTTTCATAGGAGGGGGAATGTAGTCTTATGC 2339
Qy 3121 AATACTCTTAGTCTTTGCAACATGTTAAAGTGTGTTAGCAATGCTTACAGGAGA 3180
Db 2340 AATACACTTGTAGTCTTTGCAACATGTTAAAGTGTGTTAGCAATGCTTACAGGAGA 2399
Qy 3181 GAAAGCACGCTGCATGCGATTGTTGGAAGTAAAGTGTGTTAGTGTGCTTATTAGG 3240
Db 2400 GAAAGACGACGTCATGCGATTGTTGGAAGTAAAGTGTGTTAGTGTGCTTATTAGG 2459
Qy 3241 AAGCAAACAGCGGTCTGACATGATTTGGAACCACTGAATTCGCGCATTCGAGAGAT 3300
Db 2460 AAGCAAACAGACAGTCTGACATGATTTGGAACCACTGAATTCGCGCATTCGAGAGAT 2519
Qy 3301 -ATTGTATTTAAGTGCCTAGCTCGATACAGCAACGCA-----TTTGACCATTTCCACCAT 3357
Db 2520 AATTGTATTTAAGTGCCTAGCTCGATACAGCAACGCAATTTTGAACCATTTCCACCAT 2579
Qy 3358 GGTGTGACCTT--CCAAGTTCAGCTGCGCAGCACTCAGGGCGCAAGGGCTGCTAAA 3415
Db 2580 GGTGTGACCTTCCAAAGTTCAGCTGCGCAGCACTCAGGGCGCAAGGGCTGCTAAA 2639
Qy 3416 GGAACGGAACACGCTAGAAAGCCAGTCCGCAAGAACCGTGTGATCCCGGATGAATGCA 3475
Db 2640 GGAACGGAACACGCTAGAAAGCCAGTCCGCAAGAACCGTGTGATCCCGGATGAATGCA 2699
Qy 3476 GCTACTGGGCTATCTGGAACAGGGAACCGCAAGCCGCAAGAGAGAGAGTAGCTTGCA 3535
Db 2700 GCTACTGGGCTATCTGGAACAGGGAACCGCAAGCCGCAAGAGAGAGAGTAGCTTGCA 2759
Qy 3536 GTGGCTTTACATGCGGATAGCTAGCTGGGCGGTTTTTATGACAGCAAGAGAGAGTAGCTTGCA 3595
Db 2760 GTGGGCTTTACATGCGGATAGCTAGCTGGGCGGTTTTTATGACAGCAAGAGAGTAGCTTGCA 2819
Qy 3596 TGCCAGCTGGGGCGCCCTCTGTTAAGGTGGGAAGCCCTGCAAAAGTAACTGGATGGCTT 3655
Db 2820 TGCCAGCTGGGGCGCCCTCTGTTAAGGTGGGAAGCCCTGCAAAAGTAACTGGATGGCTT 2879

QY 3656 TCTTCCGCGCAGGATCTGATGCGCGAGGGGATCAAGATCTGATCAAGACAGAGATGAG 3715
DB 2880 TCTTCCGCGCAGGATCTGATGCGCGAGGGGATCAAGATCTGATCAAGACAGAGATGAG 2939
QY 3716 GATCGTTTCCGATGATGAACAAAGATGATGACGAGGTTCTCCGCGCGCTTGGGTGG 3775
DB 2940 GATCGTTTCCGATGATGAACAAAGATGATGACGAGGTTCTCCGCGCGCTTGGGTGG 2999
QY 3776 AGAGCTATTCCGCTATGATGATGCGGCAACAGACAAATCGGCTGCTCTGATCCGCGGTGT 3835
DB 3000 AGAGCTATTCCGCTATGATGATGCGGCAACAGACAAATCGGCTGCTCTGATCCGCGGTGT 3059
QY 3836 TCCGCGCTGATGCGGCGAGGCGCGCGGTTCTTTTGTCAAGACGACCTGCTCCGCTGCC 3895
DB 3060 TCCGCGCTGATGCGGCGAGGCGCGCGGTTCTTTTGTCAAGACGACCTGCTCCGCTGCC 3119
QY 3896 TGAATGAATGACGAGGCGAGGCGCGCTATCGTGGCTGCGCCACGACGCGGCTTCCCTT 3955
DB 3120 TGAATGAATGACGAGGCGAGGCGCGCTATCGTGGCTGCGCCACGACGCGGCTTCCCTT 3179
QY 3956 GCGCAGCTGTGCTCGACCTGTGCTACTGAAGCGGGAAGGAGCTGGCTGCTATTGGGCGAAG 4015
DB 3180 GCGCAGCTGTGCTCGACCTGTGCTACTGAAGCGGGAAGGAGCTGGCTGCTATTGGGCGAAG 3239
QY 4016 TCCGCGCGCAGGATCTCCTGCTCATCTCACTTGTCTCTGCGGAGAAATGATCCATCATGG 4075
DB 3240 TCCGCGCGCAGGATCTCCTGCTCATCTCACTTGTCTCTGCGGAGAAATGATCCATCATGG 3299
QY 4076 CTGATGCAATCGCGGCGTGCATACGCTTGTGATCGGCTACCTGCCATTCGACCAACCAAG 4135
DB 3300 CTGATGCAATCGCGGCGTGCATACGCTTGTGATCGGCTACCTGCCATTCGACCAACCAAG 3359
QY 4136 GCAAACTCGCATCGAGCGAGCAGCTACTCGGATGGAAGCGGCTCTTGTGATCAGGATG 4195
DB 3360 GCAAACTCGCATCGAGCGAGCAGCTACTCGGATGGAAGCGGCTCTTGTGATCAGGATG 3419
QY 4196 ATCTGGACGAAGAGATCAGGGGCTGCGCGCAGCGGAACTGTTTCCGAGGCTCAAGGCGC 4255
DB 3420 ATCTGGACGAAGAGATCAGGGGCTGCGCGCAGCGGAACTGTTTCCGAGGCTCAAGGCGC 3479
QY 4256 GCATGCGCGAGGCGGAGATCTCGCTGATGACCCATGCGGATGCGCTTCCGATATCA 4315
DB 3480 GCATGCGCGAGGCGGAGATCTCGCTGATGACCCATGCGGATGCGCTTCCGATATCA 3539
QY 4316 TGGTGGAAAATGGCGGCTTTCTGGATTCATCGACTGTGGCGGCTGGGTGTGGCGGACC 4375
DB 3540 TGGTGGAAAATGGCGGCTTTCTGGATTCATCGACTGTGGCGGCTGGGTGTGGCGGACC 3599
QY 4376 GCTATCAGGACATAGCGTTGGCTACCCGCTGATATTGCTGAAGAGCTTGGCGCGGAATGG 4435
DB 3600 GCTATCAGGACATAGCGTTGGCTACCCGCTGATATTGCTGAAGAGCTTGGCGCGGAATGG 3659
QY 4436 CTGACCGCTTCTCGTGTCTTACGATTCGCGCTCCCGATTCGAGCGCATCGCTTCT 4495
DB 3660 CTGACCGCTTCTCGTGTCTTACGATTCGCGCTCCCGATTCGAGCGCATCGCTTCT 3719
QY 4496 ATCGCTTCTTGGACAGTTCTTCTGAGCGGAGCTCTGGGGTTCCGATA 4542
DB 3720 ATCGCTTCTTGGACAGTTCTTCTGAGCGGAGCTCTGGGGTTCCGATA 3766

RESULT 6

US-10-947-881-9
; Sequence 9, Application US/10947881
; Publication No. US20050060762A1
; GENERAL INFORMATION:
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Expression Vectors
; FILE REFERENCE: GALA-06415
; CURRENT APPLICATION NUMBER: US/10/947,881
; PRIOR FILING DATE: 2004-09-23
; PRIOR APPLICATION NUMBER: US/09/897,006

; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,851
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 5130
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-947-881-9

Query Match 30.7%; Score 1588.6; DB 9; Length 5130;
Best Local Similarity 66.1%; Pred. No. 0;
Matches 3007; Conservative 0; Mismatches 754; Indels 786; Gaps 18;

QY 1 TTTCAAGAGACCCACCCGTAGTGGCAAGCTTAAGTAACGCCACTTTGCAAGGCAT 60
DB 1 TTTGAAAGACCCACCCGTAGTGGCAAGCTTAAGTAACGCCACTTTGCAAGGCAT 60
QY 61 GGAATAATACATAACTCAGAAATAGAAAAGTTTCAGATCAAGGTCAAGGAAACAAGAAACAGC 120
DB 61 GGAATAATACATAACTCAGAAATAGAAAAGTTTCAGATCAAGGTCAAGGAAACAAGAAACAGC 120
QY 121 TGAATACAAACACAGATATCTGTGTAAAGCGGTTCTGCCCCGGCTCAGGGCCAAAGAAACA 180
DB 121 TGAATACAAACACAGATATCTGTGTAAAGCGGTTCTGCCCCGGCTCAGGGCCAAAGAAACA 180
QY 181 GATGAGCAGCTGAGTGTGATGGCCAAACAGAGATATCTGTGTAAAGCGGTTCTGCCCCGG 240
DB 181 GATGAGCAGCTGAGTGTGATGGCCAAACAGAGATATCTGTGTAAAGCGGTTCTGCCCCGG 240
QY 241 CTCGGGCGCAAGAACAGATGTTCCAGATGCGGTCAGACCTCAGCAGTTTCTAGTGAA 300
DB 241 CTCGGGCGCAAGAACAGATGTTCCAGATGCGGTCAGACCTCAGCAGTTTCTAGTGAA 300
QY 301 TCATCAGATGTTTCCAGGGTCCCAAGGACCTGAAATAGACCTGTATTTTGAAC 360
DB 301 TCATCAGATGTTTCCAGGGTCCCAAGGACCTGAAATAGACCTGTATTTTGAAC 360
QY 361 TAAACCAATCAGTTCGCTTCTGCTTCTGTCGCGGCTTCCGCTCCTCGAGCTCAATAAA 420
DB 361 TAAACCAATCAGTTCGCTTCTGCTTCTGTCGCGGCTTCCGCTCCTCGAGCTCAATAAA 420
QY 421 AGAGCCACACACCTCTCACTCGCGCGCAGCTCTTCGATAGACTGCGTCCCGGGTAC 480
DB 421 AGAGCCACACACCTCTCACTCGCGCGCAGCTCTTCGATAGACTGCGTCCCGGGTAC 480
QY 481 CCGTATTCCCAATAAAGCCTCTTGTGTCATCCGAATCGTGTCTCGCTTCTTCTT 540
DB 481 CCGTATTCCCAATAAAGCCTCTTGTGTCATCCGAATCGTGTCTCGCTTCTTCTT 540
QY 541 GGAGGGTCTCTCTGAGTGATTAACCCACAGAGAGGGGCTTTTTCATTTGGGGGCTG 600
DB 541 GGAGGGTCTCTCTGAGTGATTAACCCACAGAGAGGGGCTTTTTCATTTGGGGGCTG 600
QY 601 CCGGGATTGGAGACCCCTGCGCAGGGACACCCAGACCCACACCGGGAGGTAAAGCTGCC 660
DB 601 CCGGGATTGGAGACCCCTGCGCAGGGACACCCAGACCCACACCGGGAGGTAAAGCTGCC 660
QY 661 AGCAACTTATCTGTGTCTGTCGAGTTGTCTAGTGTCTATGTTTGAATGATGCGCTGCG 720
DB 661 AGCAACTTATCTGTGTCTGTCGAGTTGTCTAGTGTCTATGTTTGAATGATGCGCTGCG 720
QY 721 TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGGGGACCCGCTGGTGGAACTGACGAGTT 780
DB 721 TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGGGGACCCGCTGGTGGAACTGACGAGTT 780
QY 781 CTGAACACCCGCGCAACCTCGGAGACGTCACAGGACCTTTGGGGCCGCTTTTGG 840
DB 781 CTGAACACCCGCGCAACCTCGGAGACGTCACAGGACCTTTGGGGCCGCTTTTGG 840

Qy	841	CCCACCTGAGGAAGGAGTGCATGTGGAAATCCGACCCCGTCAGGATATGTGGTTCTGGT	900
Db	841	CCCACCTGAGGAAGGAGTGCATGTGGAAATCCGACCCCGTCAGGATATGTGGTTCTGGT	900
Qy	901	AGGAGCGAGAACCTAAACAGTTCCCGCCTCCGTCCTGAAATTTTTCGTTTCGGTTTCGAA	960
Db	901	AGGAGCGAGAACCTAAACAGTTTCCCGCCTCCGTCCTGAAATTTTTCGTTTCGGTTTCGAA	960
Qy	961	CCGAAGCCGCGCTTGTCTGCTGCAGCCAAAGCTTCGGCTGCAGGTGCATCTPAGAGGA	1020
Db	961	CCGAAGCCGCGCTTGTCTGCTGCAGCCAAAGCTTCGGCTGCAGGTGCAGGACTGCGGA	1020
Qy	1021	TCAATTCGGCACGAGTAAATCGGTGCTGCCGCTCTTTAGGCATATGAAGTATGSCACAGT	1080
Db	1021	CC-----CTGCACCGAAACATGGAGAAACAACATCAGGATTCCT	1059
Qy	1081	GGGATGACTTTCCTGATCAGCAAGAGACACTGCAGCTGPAACAGACTGTGTAAGTTTCG	1140
Db	1060	AGGACCCCTGCTCGTGTGTACAGCGGGGTTTTTCTTGTGTGAAGAATACTCTCAAAATACC	1119
Qy	1141	ATGCTCGCTCAGTCACAGCTTGTCTCTCCCACTCTAAATAATGGCCCAACTCTTCAAG	1200
Db	1120	ACAGAGCTTAGACTGTGGTGGACTTCTCTCAATTTTCTAGGGGAGCAACCCAGTGTCC	1179
Qy	1201	AGAGGATGAAGTCTTTATAAAACTGCACGTGATCACCTTTATCTCATTTGTGTGTAGTTC	1260
Db	1180	TGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCAACCACTCTTGTCTCTCCAAATTG	1239
Qy	1261	TCGTGCCATCATTTGGCATAGTGCAGCTCAGCTCTGAAATGGGAAACGAGAAATTTGCA	1320
Db	1240	TCCTGGGTATC-----GCTGGATGTCTCTCGCGCGTTTTTATC	1276
Qy	1321	CGGTTGGCTCAGTTAATGCAGATATATCTCCAAGTCCGGAAGGCAAAAGGAAATGGCAGTG	1380
Db	1277	ATATCTCTTTCATCTCTGCTGCTATGCTCATCTTCTTGTGTGTTCTCTGAGACTACAA	1336
Qy	1381	AAGATGAATGAGATTTTCGAAAGCTGTGATGGAACGCATGAGCAACATGGAAGCAGAA	1440
Db	1337	GGTATGTTGCCCGTTTGTCTCTACTTCCAGGA--CATCAACTACCAGCAGGACCA	1393
Qy	1441	TCCAGTATCTTTAGATAAATGAAGCCAAATCTCTAGATGCTAGAGATTTCCAAAATTTCA	1500
Db	1394	TGCAAGACCTGACAGATTCCTGCTCAAGGAACCTCTATGTTTCCCTCTT-----GTTGCT	1448
Qy	1501	GCATAACAATGATCAAGATTTTAATGATGTTCTTTTCCAGCTAAATTCCTTACTTTCCCT	1560
Db	1449	GTAACAACCTTCGGACGGAAATGCACTGTATTTCCCATCCCATCATCTCGGGCTTTTCG	1508
Qy	1561	CCATCCAGGAACATGAGAATATCATAGGGATATCTCCAAGTCATTAGTAGTCTCGAACA	1620
Db	1509	CAAGATTCCTATGGAGTGGGCTCAGTCCGTTTCTCTGG-----	1549
Qy	1621	CCACAGTACTTGATTTGCAGTTCAAGTATTTGAACAACATGAAATGGCAGAGTCCAAGAGAAATG	1680
Db	1550	-CTCAGTTTACTAGTGCCATTTGTTTCAGTGGTTC--GTAGGGCTTTCCCCCACTGTTTG	1605
Qy	1681	CATTTAAACAACAGAGAGAGATCGTAAATTTAGAGGAGCTATATACAATGCATCAGCAG	1740
Db	1606	GCTTTTCAGTTATATG-GATGATGGTATTTGGGGGCCAAGTCTGTGTAACAATCTTGAATC	1664
Qy	1741	AAATTAAGTCTCTAGATGAATAAACAAGTATATTTTGAACAGGAAATAAAAGGGGAAATGA	1800
Db	1665	CTTTTTACCTCTATTACCAATTTTCTTCTTTGGGTATACATTTTAAACCTTATAA	1724
Qy	1801	AATGTTTGAATAATATCACTAATGATCTGAGGCTGGAAGGATGGGAACAATTCACAGACAT	1860
Db	1725	AAC-----CAAAAGT	1734
Qy	1861	TGAATAATATCATTACTCCAGGTGCCAAGAGTCTCGCTGACTGGGAAATGGACCA	1920
Db	1735	TGGGGCTACTCCCTTTAACTTCATGGGATATGTAATGGATGTTGGGG-----	1781
Qy	1921	ACGATCTGGGCTCCAACATGACCATCGGGCTGTGGAACAGCAGAGGTGAAATTCACAGCCA	1980

1782	-----	1782
1981	CCTACATCAGAGCGGTAAACAGCCACATCAAAATGAGATCAAGAGTCACCACTGCATGGGA	2040
1782	-----TACTTTACCGCAAGACATATTGTACTAAAAATCAAGCAATGTTTCG-1829	
2041	CACAAACAACATCAACAGAGGACCGACCTTTTGGCTTCAACGTCAAATTTGGAAGT	2100
1830	-----AAAACTGCTGTAAATAGACCTATTGATTGGAAGATATCTCAGACACTTGTGGGT	1884
2101	TTTTCAGAGTCACCACTGTCTTACGGGCCAGTGCCTCATAGACAGGAATGGGAAGGAGG	2160
1885	CTTTTGGGCTTTGTGCCCCCTTTTACAAATGTGGCTATCTCCTGCCCTTAATGCTTTATAT	1944
2161	TCCTGAAGACCATGTGGCTCTCGGGTCAAGTGTTAATGACATTTGGTGATGACTGGAAAG	2220
1945	GCATGTATACAATCT-----	1959
2221	CTACCGGGTCGGGATCAACATCTTCACTCGCCTCGGCAACAGAGGAGTGAGTGAGTG	2280
1960	-----AAGCAGGCTTTCACCTTCTCGCAACTTACAAGGCCCTTCTCTGTGA	2005
2281	ACCAAGGTCTCTCTGGACTCCAGTGAAGGAGATAGAGGCCCTCTCGGACAAATGG	2340
2006	AACAAATATCTGAACCTTTACCCGTTGCCCGGA-----	2039
2341	TATACCAAGGCTTTCCAGGTCTAATAGGTACTCCAGGTCTTAAAGGTGATCGGGGGATCT	2400
2040	-----	2039
2401	CTGGTTTACCTGGAGTTCGAGGATTCAGGACCAATGGGAAGACCGGGAAGCCAGGAC	2460
2040	-----	2039
2461	TTAATGCAAAAGGCCAGAGGGAGAAAAAGGAGTGAAGCATGCAAGACAACTCTA	2520
2040	-----ACGCTCAG	2047
2521	ATACAGTCCAGCTGTGTGGTGGCAGCGGCCCTCACAAGGACAGAGTGGAGATTTTTCACG	2580
2048	GTCTCTGCCAAGTGTGTGCTGACGCAACCCCACTCGATGGGGCTTGGCTATCGGCATA	2107
2581	AAGGCCAGTGGGTAACGGTGTGTGACGACGGCTGGGNACTGCGTGGAGGACTGGTCGTCT	2640
2108	GCCGCATGCGCGACCTTGTGTGCTCCTCTGCGG-----	2141
2641	GCAGGAGCTGGGATACAAAGGTGTTCAAAGTGTGCATAAGCGAGCTTATTTTGGAAAAAG	2700
2142	-----ATCCNACTGCGGNACTCCTTAGCAGCTTGT-----	2171
2701	GTACGGGTCCAATATGCGTGAATGAAGTATTTTGTTCGGGAAAGAGTCATCCATTTGAAG	2760
2172	-----	2171
2761	AGTCAGAAATTAGACAGTGGGGTGTGAGACCTGTTTCGCACGACGAAGATGCTGGGGGTC	2820
2172	-----TTTGTCTCGCAGGGGCTCTGGAAGCGAA	2197
2821	ACTTTGCACCTACATAATGCATCATATTTTTCATTCACATTTTTTAAACTGTTATAAAGTG	2880
2198	ACTATCGGC-----	2207
2881	ATTTTTTTCCTTGTGCTTCACTAAATCAGCTTAATTAATTTAAGAAACTAAGAATTTT	2940
2208	-----	2207
2941	ATCCACAGAAAAGGAATATTTAAAAATCACTGGATAAACATATAAAATAGCTTCATATTT	3000
2208	-----ACCGACAACCTCT	2219
3001	GCTTCAATACCAGAACCAATTCACACTCTCTAGGTTTTTAAGTGGCTCGTCCGAATTG	3060

QY 2401 CTGGTTTACCTGGAGTTCGAGATTCCAGGACCAATGGGGAAGACCGGGAAGCCAGGAC 2460
Db 2040 ----- 2039
QY 2461 TTAATGACAAAAGGCCAGAAGGAGGAGAAAGAGGAGTGGAAGCATGCAAGACAAATCTA 2520
Db 2040 -----ACGTCAG 2047
QY 2521 ATACAGTCCGACTGGTGGGTGGCAGCGCCCTCAAGAGGAGAGTGAGATTTTTTCACG 2580
Db 2048 GTCTCTGCAAGTGTTCGTCAGCGCAACCCCACTGGATGGGGCTTGGCTATCGGCCATA 2107
QY 2581 AAGGCCAGTGGGTACGGTGTGTGACGACCCCTGGGAACCTCGGTGGAGGACTGGTCTGTCT 2640
Db 2108 GCCGCATCGCGGACCTTTGTGGCTCTCTGCG ----- 2141
QY 2641 GCAGGAGCTTGGGATACAAAGGTGTCAAAGTGTGCATAAGCGAGCTTATTTTGGMAAAG 2700
Db 2142 -----ATCCATCTCGGAACTCCTAGCAGCTTGT ----- 2171
QY 2701 GTACGGGTCCAATATATGGCTGAATGAAGTATTTTGTTCGGGAAGAGTCAATCCATTGAAG 2760
Db 2172 ----- 2171
QY 2761 AGTCAGAAATTAGACAGTGGGGTGTGAGAGCCTGTTCGACAGCAAGATGCTGGGGTC 2820
Db 2172 -----TTTGTCTCGCAGGGGCTCGAGCGAA 2197
QY 2821 ACTTTGCACCTACATAATGATCATATTTTTCATTCACATTTTAACTGTTATAAAGTG 2880
Db 2198 ACTTATCGGC ----- 2207
QY 2881 ATTTTTCCTTTCCTTCACTAAAATCAGCTTAATTAATTAATTAAGAAACTAAGAAATTT 2940
Db 2208 ----- 2207
QY 2941 ATCCACAGAAAGGAATATTTAAATAACCTGGATAAACAATAAATAGCTTCATATTT 3000
Db 2208 -----ACCGCAAACTCT 2219
QY 3001 GCTTCAATACCAGAACCAATTTCACTTCTAGGTCTTAAAGTGGTCTGTCGGAATG 3060
Db 2220 GTTGTCTCTCTCGGAATAACACTCTCTTCCATGGCTGTAGGGTGTGTCGCAACTGG 2279
QY 3061 ATCCCTCTCAGGATATAGTAGTTTCGCTTTTCATAGGAGGGGAAATGTAGTCTTATGC 3120
Db 2280 ATCCCTCTCAGGATATAGTAGTTTCGCTTTTCATAGGAGGGGAAATGTAGTCTTATGC 2339
QY 3121 AATACTCTTGTAGTCTTTCGCAACATGGTAAAGATAGTTCAGCAATGCAATGCTTTACAGGAGA 3180
Db 2340 AATACTCTTGTAGTCTTTCGCAACATGGTAAAGATAGTTCAGCAATGCAATGCTTTACAGGAGA 2399
QY 3181 GAAAAGCAGCGTCATGCCGATTTGGTGGAGTAAAGTGGTACCATGCTGCTTATTAGG 3240
Db 2400 GAAAAGCAGCGTCATGCCGATTTGGTGGAGTAAAGTGGTACCATGCTGCTTATTAGG 2459
QY 3241 AAGGCAACAGACGGGTCTGACATGGATTGGACGAACCACTGAAATTCGCAATGCGAGAT 3300
Db 2460 AAGGCAACAGACAGGTCTGACATGGATTGGACGAACCACTGAAATTCGCAATGCGAGAT 2519
QY 3301 -ATTGTATTTAAGTGCCTAGTCTGATACAGCAACGCCA --TTTGACCAATTCACCAAT 3357
Db 2520 AATTGTATTTAAGTGCCTAGTCTGATACAGCAACGCCAATTTTGAACCAATTCACCAAT 2579
QY 3358 GGTGTGACCTT--CCAGCTTTCAGCTGCGCGCAAGCACTCAGGGCGCAAGGGCTGCTAAA 3415
Db 2580 GGTGTGACCTTTCAAAAGCTTTCAGCTGCGCGCAAGCACTCAGGGCGCAAGGGCTGCTAAA 2639
QY 3416 GGAAGCGGAACACCTGAGAAAGCAGTCCGAGAAACCGTGTGACCCCGGATGAATGTCA 3475
Db 2640 GGAAGCGGAACACCTGAGAAAGCAGTCCGAGAAACCGTGTGACCCCGGATGAATGTCA 2699

QY 3476 GCTACTGGGCTATCTGCAAGGGAACCGACCGCAAGAGAAAGACAGAGTAGCTTCCA 3535
Db 2700 GCTACTGGGCTATCTGACCAAGGGGAAACCGCAAGCGCAAGAGAAAGACAGAGTAGCTTCCA 2759
QY 3536 GTGGGCTTTACATGCGGATAGCTAGACTGGGCGGTTTTATGACAGCAAGCAAGCAACCGGAAT 3595
Db 2760 GTGGGCTTTACATGCGGATAGCTAGACTGGGCGGTTTTATGACAGCAAGCAAGCAACCGGAAT 2819
QY 3596 TGCCAGCTGGGGCCCTCTCTGTTAAGGTTGGGAAGCCCTGCAAAAGTAAACTGGATGGCTT 3655
Db 2820 TGCCAGCTGGGGCCCTCTCTGTTAAGGTTGGGAAGCCCTGCAAAAGTAAACTGGATGGCTT 2879
QY 3656 TCTTGGCCCAAGAGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGACAGAGATGAG 3715
Db 2880 TCTTGGCCCAAGAGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGACAGAGATGAG 2939
QY 3716 GATCGTTTCGATCATGATTGAACAAGATGATTGACAGCAGAGTTCTCCGCGCTTGGGTGG 3775
Db 2940 GATCGTTTCGATCATGATTGAACAAGATGATTGACAGCAGAGTTCTCCGCGCTTGGGTGG 2999
QY 3776 AGAGGCTTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCCGCTGT 3835
Db 3000 AGAGGCTTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCCGCTGT 3059
QY 3836 TCCGCTGTGACGCGCAGGGGCGCCGGTTCTTTTGTCAAGACCGACTCTGCTGCGGTGCC 3895
Db 3060 TCCGCTGTGACGCGCAGGGGCGCCGGTTCTTTTGTCAAGACCGACTCTGCTGCGGTGCC 3119
QY 3896 TGAATGAATGACAGCAGGAGCGCGGCTATCGTGGCTGGCCACGACCGGCGTTCCTT 3955
Db 3120 TGAATGAATGACAGCAGGAGCGCGGCTATCGTGGCTGGCCACGACCGGCGTTCCTT 3179
QY 3956 GCGCAGCTGTCTCGAGCTTGTCACTGAAGCGGGAAGGAGTGTGCTGCTATTGGGCGAAG 4015
Db 3180 GCGCAGCTGTCTCGAGCTTGTCACTGAAGCGGGAAGGAGTGTGCTGCTATTGGGCGAAG 3239
QY 4016 TGCCGGGCGAGGATCTCTGTCACTCACTTGTCTGCTGCGGAGAAAGTATCCATCATGG 4075
Db 3240 TGCCGGGCGAGGATCTCTGTCACTCACTTGTCTGCTGCGGAGAAAGTATCCATCATGG 3299
QY 4076 CTGATGAATGCGGCGGCTGCATACGCTTGTATCGGCTACTCGGCTTCCGACCCACCAAG 4135
Db 3300 CTGATGAATGCGGCGGCTGCATACGCTTGTATCGGCTACTCGGCTTCCGACCCACCAAG 3359
QY 4136 CGAAACATCGCATCGAGCAGCAGTACTCGGATGGAAGCGGCTTGTGCTCATCAGGATG 4195
Db 3360 CGAAACATCGCATCGAGCAGCAGTACTCGGATGGAAGCGGCTTGTGCTCATCAGGATG 3419
QY 4196 ATCTGACGAAGAGCATCAGGGGCTCGCGCAGCGCAACTGTTTCGCGAGGCTCAAGGCGC 4255
Db 3420 ATCTGACGAAGAGCATCAGGGGCTCGCGCAGCGCAACTGTTTCGCGAGGCTCAAGGCGC 3479
QY 4256 GCATGCCCGAGCGGAGGATCTCGTGTGACCCATGCGGATGCTGCTGCGGAATATCA 4315
Db 3480 GCATGCCCGAGCGGAGGATCTCGTGTGACCCATGCGGATGCTGCTGCGGAATATCA 3539
QY 4316 TGGTGGAAAAATGGCCGCTTTTCTGGATTTCATCGACTGTGCGCGGCTGGGTGTCGCGAACC 4375
Db 3540 TGGTGGAAAAATGGCCGCTTTTCTGGATTTCATCGACTGTGCGCGGCTGGGTGTCGCGAACC 3599
QY 4376 GCTATCAGGACATAGCGTTCGGCTACCGTGAATTTGCTGAAGAGCTTTGGCGGCGAATGGG 4435
Db 3600 GCTATCAGGACATAGCGTTCGGCTACCGTGAATTTGCTGAAGAGCTTTGGCGGCGAATGGG 3659
QY 4436 CTGACCGCTTCTCGTCTTTTACGGTATCCGCGCTCCCGATTCGACGGCATCGCCTTCT 4495
Db 3660 CTGACCGCTTCTCGTCTTTTACGGTATCCGCGCTCCCGATTCGACGGCATCGCCTTCT 3719
QY 4496 ATCGCTTCTTGAAGAGTCTCTCTGAGCGGAGCTCTGGGGTTCGATA 4542
Db 3720 ATCGCTTCTTGAAGAGTCTCTCTGAGCGGAGCTCTGGGGTTCGATA 3766

RESULT 8
US-11-036-557-9
; Sequence 9, Application US/11036557
; Publication No. US20050221429A1
; GENERAL INFORMATION:
; APPLICANT: Black, Gregory T.
; TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors Comprising an
; FILE REFERENCE: GALA-09479
; CURRENT APPLICATION NUMBER: US/11/036,557
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9
; LENGTH: 5130
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-11-036-557-9

Query Match 30.7%; Score 1588.6; DB 10; Length 5130;
Best Local Similarity 66.1%; Pred. No. 0;
Matches 3007; Conservative 0; Mismatches 754; Indels 786; Gaps 18;

Qy 1 TTTGAAGACCCACCGTAGTGGCAAGCTTAGCTTAAGTAACGCCACTTTTGCAGGCAT 60
Db 1 TTTGAAGACCCACCGTAGTGGCAAGCTTAGCTTAAGTAACGCCACTTTTGCAGGCAT 60

Qy 61 GGAAAAATACATACTAGAAATAGAAAAGTTTCAGATCAAGGTCAGGAAACAAAGAACAGC 120
Db 61 GGAAAAATACATACTAGAAATAGAAAAGTTTCAGATCAAGGTCAGGAAACAAAGAACAGC 120

Qy 121 TGAATACCAACAGGATATCTGTGTAGCCGGTTCCTGCCCGGCTCAGGCCCAAGAAC 180
Db 121 TGAATACCAACAGGATATCTGTGTAGCCGGTTCCTGCCCGGCTCAGGCCCAAGAAC 180

Qy 181 GATGAGACAGCTGAGTGTAGTGGCCAAACAGGATATCTGTGTAGGAGCTTCTCCCGCG 240
Db 181 GATGAGACAGCTGAGTGTAGTGGCCAAACAGGATATCTGTGTAGGAGCTTCTCCCGCG 240

Qy 241 CTCGGGGCCAAAGACAGATGGTCCCAGATCGGTCAGCCCTCAGCAGTTTCTAGTGAA 300
Db 241 CTCGGGGCCAAAGACAGATGGTCCCAGATCGGTCAGCCCTCAGCAGTTTCTAGTGAA 300

Qy 301 TCATCAGATGTTTCCAGGGTGCCTCCAGGACCTGAAATGACCTGTACCTTATTTGAAC 360
Db 301 TCATCAGATGTTTCCAGGGTGCCTCCAGGACCTGAAATGACCTGTACCTTATTTGAAC 360

Qy 361 TAACCAATCAGTTCGCTTCTCGCTTCTGTTCGGCGCTTCCGCTCTCCGAGCTCAATAA 420
Db 361 TAACCAATCAGTTCGCTTCTCGCTTCTGTTCGGCGCTTCCGCTCTCCGAGCTCAATAA 420

Qy 421 AGAGCCCAACACCCCTCACTCGGGCGCCAGTCTTCCGATAGACTGGCTGCCCGGGTAC 480
Db 421 AGAGCCCAACACCCCTCACTCGGGCGCCAGTCTTCCGATAGACTGGCTGCCCGGGTAC 480

Qy 481 CCGTATTTCCCAATAAAGCTTCTGTGTTCGATCCGATCCGATCGTGTCTCGCTTCTCT 540
Db 481 CCGTATTTCCCAATAAAGCTTCTGTGTTCGATCCGATCCGATCGTGTCTCGCTTCTCT 540

Qy 541 GGAGGGTCTCTCTGAGTGATTGACTACCCAGAGCGGGGTCTTTTCAATTTGGGGGCTCGT 600
Db 541 GGAGGGTCTCTCTGAGTGATTGACTACCCAGAGCGGGGTCTTTTCAATTTGGGGGCTCGT 600

Qy 601 CCGGGATTTGGAGACCCCTGCGCCAGGACCAACCCAGCCAGGAGGTAAAGCTGGCC 660
Db 601 CCGGGATTTGGAGACCCCTGCGCCAGGACCAACCCAGCCAGGAGGTAAAGCTGGCC 660

Qy 661 AGCAACTTATCTGTGTCTGTCCGATGTCTAGTGTCTATGTTTATGTCGCGCTGGG 720
Db 661 AGCAACTTATCTGTGTCTGTCCGATGTCTAGTGTCTATGTTTATGTCGCGCTGGG 720

Qy 721 TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCCGCTGGTGGAACTGACGAGTT 780
Db 721 TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCCGCTGGTGGAACTGACGAGTT 780

Qy 781 CTGAACACCCCGCGCAACCTCTGGAGACGCTCCAGGAGCTTTGGGGGCGCTTTTGTGG 840
Db 781 CTGAACACCCCGCGCAACCTCTGGAGACGCTCCAGGAGCTTTGGGGGCGCTTTTGTGG 840

Qy 841 CCGGACCTGAGGAGGAGTGGATGCGAATCCGACCCCGCTCAGGATATGTTGTTCTGGT 900
Db 841 CCGGACCTGAGGAGGAGTGGATGCGAATCCGACCCCGCTCAGGATATGTTGTTCTGGT 900

Qy 901 AGGAGACGAGAACCTAAACAGATTCCCGCTCTGAAATTTTGTCTTTCGGTTTGGAA 960
Db 901 AGGAGACGAGAACCTAAACAGATTCCCGCTCTGAAATTTTGTCTTTCGGTTTGGAA 960

Qy 961 CCGAAGCCCGCGCTTGTCTGTGACCAAGCTTGGGCTGACAGTCTGAGGAGCTGGGA 1020
Db 961 CCGAAGCCCGCGCTTGTCTGTGACCAAGCTTGGGCTGACAGTCTGAGGAGCTGGGA 1020

Qy 1021 TCAATTCGGCAGGATTAATCGGCTGCTGCTTTTAGGACATATGAAGTATGACACAGT 1080
Db 1021 CC-----CTGCACCGAACATGGAGAACACACATCAGGATTCCT 1059

Qy 1081 GGGATGACTTTCTCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTGC 1140
Db 1060 AGGACCCCTGCTGCTGTGTACAGGGGGTCTTCTGTTGACAGAGATCTCACAATACC 1119

Qy 1141 ATGCTGCTCAGTGACAGCTTTGCTTCTCCCATCTTAAATGGCCCAACTCTTCAAG 1200
Db 1120 ACAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAGCACCCAGTGTCC 1179

Qy 1201 AGAGGATGAGTCTTATAAATGCACTGATCACTTATCTCATTTGTTTGTAGTTTC 1260
Db 1180 TGGCAAAATTCGAGTCCCAACCTCCAACTACTCAACCTCTTGTCTCTCAATTTG 1239

Qy 1261 TCGTGGCCCATCAITGGCATAGTGGCAGCTCAGCTCTCTGAAATGGGAAACGAAGATTCGA 1320
Db 1240 TCTGGGTATC-----GCTGGATGTCTGCGGCGTTTATC 1276

Qy 1321 CGGTGGCTCAGTTAATGACAGATATATCTCAAGTCCGGAAGGCAAGAAATGGCAGTG 1380
Db 1277 ATATTCTCTTCATCTCTGCTGCTATGCTCATCTCTTGTGTTGTTCTTCTGGACTACCA 1336

Qy 1381 AAGTGAATGAGATTTTCGAGAGCTGTGTGAGCCATGAGCAACATGGAAGACGAA 1440
Db 1337 GGTATGTTGCCGCTTGTCTCTACTTCCAGGAA---CATCAACTACAGCACGGGACCA 1393

Qy 1441 TCCAGTATCTTTTCAGATAATGAAGCCCAATCTCTAGATGCTAAGAAATTCGAAATTTCA 1500
Db 1394 TGCAGACCTGACAGATCTCTCTCAAGGAACCTCTATGTTCCCTCTT-----GTTGCT 1448

Qy 1501 GCATAACAACCTGATCAAGATTTAATGATGTTCTTTTCCAGCTAAATTTCTTACTTCTCT 1560
Db 1449 GTACAAAACCTTCGGACGGAAACCTGCACTTGTATTCCTCATCCCATCTCTGGGCTTCG 1508

Qy 1561 CCATCCAGGACATGAGAAATATCATAGGGGATATCTCCAAGTCAATAGTAGTCTGAACA 1620
Db 1509 CAAGATTCCTATGGGAGTGGGCTCAGTCCGTTTCTCTCTG----- 1549

Qy 1621 CCAGAGTACTTGTATTTGACGTTTCAATTTGAACACATGAAATGGCAGAGTCCAGAGAAATG 1680
Db 1550 -CTCAGTTTACTAGTGCCATTTGTTTCAAGTGGTTC---GTAGGGCTTTTCCCGACTGTTT 1605

Qy 1681 CATTTAAACAACAGAGGAGATGCGTAAATTAGAGGAGCGTATATACAAATGCATCAGAG 1740
Db 1606 GCITTCAGTTATATG-GATGATGTTGTTTGGGGCCCAAGTCTGTACAACTCTTGAGTC 1664

Qy 1741 AATTAAGTCTCTAGATGAAGAAACAGTATATTTGGAACAGGAAATAAAGGGGAATGA 1800
Db 1665 CTTTTTACCTCTATTAACCAATTTTCTTGTCTTGTGGGTATATCAATTTAAACCCCTATAA 1724

Qy 1801 AACTGTTGAATAATATCACTAATGATCTGAGGCTGGAAGGATTTGGGAACATTTCTCAGACAT 1860

Db 1725 AAC ||||| 1734
Qy 1861 TGAATAATACATTTACTCAAGTGCAGAAAGTCTCCTGACTGGGAAATGGACCA 1920
Db 1735 TGGGCTACTCCCTTAACCTTATGGATATGTAATGGATGTTGGG 1781
Qy 1921 ACGATCTGGGCTCCAAATGACATGACCATCGGGGCTGTGAACAGCAGAGGTGAATTCACAGGCA 1980
Db 1782 1781
Qy 1981 CCTACATCACAGCGGTAAACAGCCACATCAATGAGATCAAGAGTCAACACTGCATGGGA 2040
Db 1782 TACTTTACCCCAAGAACATATGTACTAAATCAAGCAATGTTTCG- 1829
Qy 2041 CACAAAACACCATCAACAAAGAGGACCCAGGCCACTTTGGCTTCAACGCTCAATTTGGAAGT 2100
Db 1830 AAAAGCTCTGTAATAGACCTATTGATTTGGAAGTATGTAGAGACTTTGGGT 1884
Qy 2101 TTTGAGATCCACACTGTCTTTCAGGGGCCAGTGTCTTATAGACAGGAATGGGAAGG 2160
Db 1885 CTTTGGGCTTTGTGCCCCCTTTTACACAAATGTGGCTATCCTGCTTAATGCCCTTTATAT 1944
Qy 2161 TCCTGAGACCATGTGGCTGTGCGGTCAAGTGTAAATGACATTTGGTGTGACTGGAAAG 2220
Db 1945 GCATGTATACAACT- 1959
Qy 2221 CTACAGGGTGGGATCAACATCTTCACTCGCTGCGCACACAGAAAGAGTGAAGTGT 2280
Db 1960 AAGAGGCTTTCACTTTCTCGCCAACTTACAGGCCCTTTCTGTGTA 2005
Qy 2281 ACCAAGTCTCTGGAATCTCAGGTGAAAAGGAGATAGAGCCCTCTGTGCAAAAATGG 2340
Db 2006 AACAAATCTGAACCTTTACCCCGTTGCCCGCA 2039
Qy 2341 TATACAGGCTTTCCAGTCTTAATAGTACTCCAGGTCTTAAGGTGATCGGGGGATCT 2400
Db 2040 2039
Qy 2401 CTGGTTTACCTGGAGTTCGAGGATTCGCCAGGACCAATGGGAAGACCGGAAGCCAGGAC 2460
Db 2040 2039
Qy 2461 TTAATGACAAAAAGGCCAGAAAGGAGGAGTGAAGCATGCMAAGACAAATCTA 2520
Db 2040 ACGGTCAG 2047
Qy 2521 ATACAGTCCGACTGGTGGTGGCAGCGGCCCTCAGAGGAGAGTGGAGATTTTTCAG 2580
Db 2048 GTCTCTGCAAGTGTGTGTCAGCGCAACCCCACTGGATGGGGCTTGGCTATCGGCCATA 2107
Qy 2581 AAGCCAGTGGGTACGGTGTGTCAGCACCGCTCGGAACCTCGTGGAGGACTGGTCTCT 2640
Db 2108 CCGCATGCGCGGACCTTTGTGGCTCTCTGCGG- 2141
Qy 2641 GCAGAGCTTGGGATACAAAGGTGTTCAAAGTGTGTCATAGCGAGCTTAATTTGGAAG 2700
Db 2142 ATCCATCTGCGGAACCTCTAGCAGCTTGT- 2171
Qy 2701 GTACGGGTCCAAATATGGCTGAATGAATTTGTTTCGGGAAGAGTCATCCATTTGAAG 2760
Db 2172 2171
Qy 2761 AGTGCAGAAATTAGACAGTGGGTGTGAGAGCTTGTGACAGCAGAGATGCTGGGGTC 2820
Db 2172 TTTGCTCGCAGCGGCTCGGAGCGAA 2197
Qy 2821 ACTTTGACCTTACATAATATGATCATATTTTCAATTCACATTTTAACTGTTATAAGTG 2880
Db 2198 ACTTATCGG- 2207
Qy 2881 ATTTTTCCTTGTGCTTCACTAAATACAGTTAATTAATTAATTTAAGAACTAAGAAATTTT 2940

Db 2208 2207
Qy 2941 ATCCACAGAAAGGAATATTTAAAAATCACTGGATAAACATATAAAATAGCTTCATATTT 3000
Db 2208 ACCGACAACTCT 2219
Qy 3001 GCTTCAATACCAGAACCAATTTCAACTTCTCTAGTGTTTTAAAGTGGCTCGTGGCGAATTG 3060
Db 2220 GTTGTCTCTCTCGGAATAACACCTCTCTTCCATGGCTGTAGGGTGTGCTGCCAACTGG 2279
Qy 3061 ATCCCTCTCAGATATAGTATGTTTCGCTTTTGCATAGGGAGGGGAAATGTAGTCTTATGC 3120
Db 2280 ATCCCTCTCAGATATAGTATGTTTCGCTTTTGCATAGGGAGGGGAAATGTAGTCTTATGC 2339
Qy 3121 AATACTCTTGTAGTCTTTCACCATGTAACGATGAGTTAGCAACATGCTTTTACAAGGAGA 3180
Db 2340 AATACTCTTGTAGTCTTTCACCATGTAACGATGAGTTAGCAACATGCTTTTACAAGGAGA 2399
Qy 3181 GAAAAGCACCGTGCATGCGGATTTGGTGGAAAGTAAAGTGGTACGATCGTGCCTTATTAGG 3240
Db 2400 GAAAAGCACCGTGCATGCGGATTTGGTGGAAAGTAAAGTGGTACGATCGTGCCTTATTAGG 2459
Qy 3241 AAGSAAACAGACGGTCTGATGATGATTTGGACGAACCACTGAATTCGSCATTTGACAGAT 3300
Db 2460 AAGSAAACAGACGGTCTGATGATGATTTGGACGAACCACTGAATTCGSCATTTGACAGAT 2519
Qy 3301 -ATTGTATTTAAGTGCCTAGCTCGATACAGCAACGCCA--TTTGACCATTTTACCACATTT 3357
Db 2520 AATTGTATTTAAGTGCCTAGCTCGATACAGCAACGCCAATTTTGAACATTTACCACATTT 2579
Qy 3358 GGTGTGCACCT--CCAAAGCTTTCAGCTGCGCGCAAGCACTCAGGGCGGAGGGCTGTAAA 3415
Db 2580 GGTGTGCACCTTCCAAAGCTTTCAGCTGCGCGCAAGCACTCAGGGCGGAGGGCTGTAAA 2639
Qy 3416 GGAAGCGGAACACGTAGAAAGCCAGTCCGAGAAACCGTCTGACCCCGGATGAATGTCA 3475
Db 2640 GGAAGCGGAACACGTAGAAAGCCAGTCCGAGAAACCGTCTGACCCCGGATGAATGTCA 2699
Qy 3476 GCTACTGGGCTATCTGACCAAGGGAACCGCAAGCAAGAGAGAGAGAGAGAGAGAGAGAG 3535
Db 2700 GCTACTGGGCTATCTGACCAAGGGAACCGCAAGCGCAAGAGAGAGAGAGAGAGAGAG 2759
Qy 3536 GTGGGCTTACATGCGGATAGCTAGACTGGGCGGTTTTATGACACAGCAAGCGGAAT 3595
Db 2760 GTGGGCTTACATGCGGATAGCTAGACTGGGCGGTTTTATGACACAGCAAGCGGAAT 2819
Qy 3596 TGCCAGCTGGGGCGCCCTCTGTTAAGGTTGGGAAGCCCTGCAAAAGTAACTGGATGGCTT 3655
Db 2820 TGCCAGCTGGGGCGCCCTCTGTTAAGGTTGGGAAGCCCTGCAAAAGTAACTGGATGGCTT 2879
Qy 3656 TCTTCCGCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAG 3715
Db 2880 TCTTCCGCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAG 2939
Qy 3716 GATCGTTTCGATGATTTGAACAGATGGAATGACAGGATTTCTCCGGCCGCTTGGGTGG 3775
Db 2940 GATCGTTTCGATGATTTGAACAGATGGAATGACAGGATTTCTCCGGCCGCTTGGGTGG 2999
Qy 3776 AGAGGCTATTGCGCTATGACTGGGCAACAGCAATCGGCTGCTGTGATGCGCCCGCTGT 3835
Db 3000 AGAGGCTATTGCGCTATGACTGGGCAACAGCAATCGGCTGCTGTGATGCGCCCGCTGT 3059
Qy 3836 TCCGGCTGTGACGCGAGGGCGCCCGGTTCTTTTGTCAAGACCGACTGCTGCCGTGCC 3895
Db 3060 TCCGGCTGTGACGCGAGGGCGCCCGGTTCTTTTGTCAAGACCGACTGCTGCCGTGCC 3119
Qy 3896 TGAATGAATCTCAGGACGAGCGCGCTTATCGTGGCTGGCCACGACGCGGCTTCTT 3955
Db 3120 TGAATGAATCTCAGGACGAGCGCGCTTATCGTGGCTGGCCACGACGCGGCTTCTT 3179
Qy 3956 GCGCAGCTGTGCTGACGCTTGTCTGACGCGGAGAGGACTGCGCTGCTATTGGGCGAG 4015
Db 3180 GCGCAGCTGTGCTGACGCTTGTCTGACGCGGAGAGGACTGCGCTGCTATTGGGCGAG 3239

4016 TGCCGGGCGAGGATCTCTGTGTCATCTCACCTTGCTCTCCGCGAGAAAGTATCCATCATGG 4075
Db TGCCGGGCGAGGATCTCTGTGTCATCTCACCTTGCTCTCCGCGAGAAAGTATCCATCATGG 3299
4076 CTGATGCAATGCGGGCTGTCATACGCTTGATCGGGCTACTGCGCCATTCGACCAACCAAG 4135
Db CTGATGCAATGCGGGCTGTCATACGCTTGATCGGGCTACTGCGCCATTCGACCAACCAAG 3359
4136 CGAAACATCGCATCGAGCGAGCAGCTACTCGGATGGAAGCCGGCTTGTCTCATCAGGATG 4195
Db CGAAACATCGCATCGAGCGAGCAGCTACTCGGATGGAAGCCGGCTTGTCTCATCAGGATG 3419
4196 ATCTGGACGAAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGGCG 4255
Db ATCTGGACGAAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGGCG 3479
4256 GCATGCCCGAGGAGGATCGTCTGATGACCCATGCGCGATGCCCTGTTCGCCGAATATCA 4315
Db GCATGCCCGAGGAGGATCGTCTGATGACCCATGCGCGATGCCCTGTTCGCCGAATATCA 3539
4316 TGGTGGAAATGCGCGCTTTCTGATTCATCGACTGTGGCGGCTGGGTGTGGCGGACC 4375
Db TGGTGGAAATGCGCGCTTTCTGATTCATCGACTGTGGCGGCTGGGTGTGGCGGACC 3599
4376 GCTATCAGGACATAGCGTTGGCTACCCGCTGATATGCTGAAGAGCTTGGCGGCGAATGGG 4435
Db GCTATCAGGACATAGCGTTGGCTACCCGCTGATATGCTGAAGAGCTTGGCGGCGAATGGG 3659
4436 CTGACCGCTTCTCTGCTTTACGGTATCGCGCTCCCGATTCGCGAGCGCATCGCTTCT 4495
Db CTGACCGCTTCTCTGCTTTACGGTATCGCGCTCCCGATTCGCGAGCGCATCGCTTCT 3719
4496 ATCGGCTTCTGACGAGTTCTTCTGAGCGGAGCTCTGGGGTTCGATA 4542
Db ATCGGCTTCTTACGAGTTCTTCTGAGCGGAGCTCTGGGGTTCGATA 3766

RESULT 9

US-08-786-531B-3
; Sequence 3, Application US/08786531B
; Publication No. US20020015979A1
; GENERAL INFORMATION:
; APPLICANT: Link, Charles J.
; APPLICANT: Levy, John P.
; APPLICANT: Wang, Suming
; APPLICANT: Seregina, Tatiana
; TITLE OF INVENTION: Vehicles for Stable Transfer of Green
; TITLE OF INVENTION: Fluorescent Protein Gene and Methods of Use for Same
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
; STREET: 801 Grand Suite 3200
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/786,531B
; FILING DATE: 21-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/010371
; FILING DATE: 22-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Nebel, Heidi S.
; REGISTRATION NUMBER: 37,719
; REFERENCE/DOCKET NUMBER: hgtri

TELECOMMUNICATION INFORMATION:

; TELEPHONE: 515-288-3667
; TELEFAX: 515-288-1338
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6620 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-786-531B-3

Query Match 28.7%; Score 1486; DB 2; Length 6620;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3692 GATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAAACAAGATGGATTGCACG 3751
Db 2777 GATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAAACAAGATGGATTGCACG 2836
Qy 3752 CAGGTTCTCCGGCGCTTGGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAAGACAA 3811
Db 2837 CAGGTTCTCCGGCGCTTGGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAAGACAA 2896
Qy 3812 TCGGCTGCTGTGATGCGCGCTGTTCCGGCTGTCCAGCGCAGGGGCGCCGGTCTTTTGG 3871
Db 2897 TCGGCTGCTGTGATGCGCGCTGTTCCGGCTGTCCAGCGCAGGGGCGCCGGTCTTTTGG 2956
Qy 3872 TCAAGACCGACCTGTCTCGGTGCCCTGAATGAATGCAAGACGAGGACGAGCGCGGTATCGT 3931
Db 2957 TCAAGACCGACCTGTCTCGGTGCCCTGAATGAATGCAAGACGAGGACGAGCGCGGTATCGT 3016
Qy 3932 GGCTGGCCACGACGCGGCTTCTTTCGCGAGCTGTGCTCGAGCTGTGCTCACTGAAGCGGAA 3991
Db 3017 GGCTGGCCACGACGCGGCTTCTTTCGCGAGCTGTGCTCGAGCTGTGCTCACTGAAGCGGAA 3076
Qy 3992 GGAGCTGGCTGCTATTGGGCGAAGTGGCGGCGAGGATCTCTGTGATCTCACCTTGCTC 4051
Db 3077 GGAGCTGGCTGCTATTGGGCGAAGTGGCGGCGAGGATCTCTGTGATCTCACCTTGCTC 3136
Qy 4052 CTGCCGAGAAAGTATCCATCATGCTGATGCAATGCGCGCGCTGCATACGCTTGCATCCGG 4111
Db 3137 CTGCCGAGAAAGTATCCATCATGCTGATGCAATGCGCGCGCTGCATACGCTTGCATCCGG 3196
Qy 4112 CTACCTGCCATTCGACCAACAAGCGAAACATGCGATCGAGCGAGCAGTACTCGGATGG 4171
Db 3197 CTACCTGCCATTCGACCAACAAGCGAAACATGCGATCGAGCGAGCAGTACTCGGATGG 3256
Qy 4172 AAGCGGCTTGTGTCGATCAGGATGATCTGGAACGAGAGCATCAGGGGCTCGCGCCAGCG 4231
Db 3257 AAGCGGCTTGTGTCGATCAGGATGATCTGGAACGAGAGCATCAGGGGCTCGCGCCAGCG 3316
Qy 4232 AACTGTTTCGCCAGGCTCAAGCGCGCATGCGCCGAGGAGTCTCGTGTGACCCCATG 4291
Db 3317 AACTGTTTCGCCAGGCTCAAGCGCGCATGCGCCGAGGAGTCTCGTGTGACCCCATG 3376
Qy 4292 GCGATGCTGCTTCCCGCAATATCATGTTGAAATATGCGCGCTTTTCTGGAATTCATCGACT 4351
Db 3377 GCGATGCTGCTTCCCGCAATATCATGTTGAAATATGCGCGCTTTTCTGGAATTCATCGACT 3436
Qy 4352 GTGCCGCGCTGGGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGGTGATATTG 4411
Db 3437 GTGCCGCGCTGGGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGGTGATATTG 3496
Qy 4412 CTGAAGAGCTTGGCGGCAATGGCTGACCGCTTCTCTGCTGCTTTACGGTATCGCCGCTC 4471
Db 3497 CTGAAGAGCTTGGCGGCAATGGCTGACCGCTTCTCTGCTGCTTTACGGTATCGCCGCTC 3556
Qy 4472 CCGATTGCGAGCGCATCGCCTTCTATCGCCTTCTTGAACGAGTTCTTCTGAGCGGAGTCT 4531
Db 3557 CCGATTGCGAGCGCATCGCCTTCTATCGCCTTCTTGAACGAGTTCTTCTGAGCGGAGTCT 3616

Db 3348 AGAACAGATGGTCCCGCAGATGGGTCCAGCCCTCAGCAGATTCTTAGAGAACCATCAGATG 3407
Qy TTTCCAGGGTCCCGCAGGACCTGAATGACCCCTGTGCTTATTGTGAACCTAACCAATCAG 4956
Db 3408 TTTCCAGGGTCCCGCAGGACCTGAATGACCCCTGTGCTTATTGTGAACCTAACCAATCAG 3467
Qy TTCGCTTCTCGCTTCTGTTCGGCGCTTCTGCTCCCGAGCTCAATAAAGAGGCCACAA 5016
Db 3468 TTCGCTTCTCGCTTCTGTTCGGCGCTTCTGCTCCCGAGCTCAATAAAGAGGCCACAA 3527
Qy 5017 CCCCTCACTCGGGCGCAGCTCTCCGATTGACTGAGTCGCGCGGTACCCGCTGATCCA 5076
Db 3528 CCCCTCACTCGGGCGCAGCTCTCCGATTGACTGAGTCGCGCGGTACCCGCTGATCCA 3587
Qy 5077 ATAAACCTCTTGCAGTTGCAATCCGACTTGCTGCTGCTTCTTCTTGGAGGGTCTCCT 5136
Db 3588 ATAAACCTCTTGCAGTTGCAATCCGACTTGCTGCTGCTTCTTCTTGGAGGGTCTCCT 3647
Qy 5137 CTGAGTGATTGACTACCCGTCAGCGGGGTCTTTCAATTTGG 5177
Db 3648 CTGAGTGATTGACTACCCGTCAGCGGGGTCTTTCAATTTGG 3688

RESULT 11

US-10-987-388-36

; Sequence 36, Application US/10987388

; Publication No. US2005015871A1

; GENERAL INFORMATION:

; APPLICANT: Stratagene California

; APPLICANT: Braman, Jeffrey

; APPLICANT: Carstens, Carsten-Peter

; APPLICANT: Novoradovskaya, Natalia

; APPLICANT: Bagga, Rajesh

; APPLICANT: Baschore, Lee Scott

; TITLE OF INVENTION: Compositions and Methods for Protein Isolation

; FILE REFERENCE: 25436/2465

; CURRENT APPLICATION NUMBER: US/10/987,388

; CURRENT FILING DATE: 2004-11-12

; PRIOR APPLICATION NUMBER: US 10/712,137

; PRIOR FILING DATE: 2003-11-13

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 36

; LENGTH: 6825

; TYPE: DNA

; ORGANISM: Artificial sequence

; FEATURE:

; OTHER INFORMATION: Retroviral vector pFB-CTAP-nso

US-10-987-388-36

Query Match 28.3%; Score 1463; DB 9; Length 6825;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3715 GGATCGTTTCCCATGATTGAACAAGATGGATTGACGCGAGTCTCCGGCGCGCTTGGGTG 3774
Db 3005 GGATCGTTTCCCATGATTGAACAAGATGGATTGACGCGAGTCTCCGGCGCGCTTGGGTG 3064
Qy 3775 GAGAGGCTATTCCGCTATGACTGGGCAACAAGACAATCGGCTGCTCTGATGCCCGCGTG 3834
Db 3065 GAGAGGCTATTCCGCTATGACTGGGCAACAAGACAATCGGCTGCTCTGATGCCCGCGTG 3124
Qy 3835 TTCGGCTGTGACGCGAGGGCGCGCTTCTTTTGTCAAGACCGACTGTCGGGTGCC 3894
Db 3125 TTCGGCTGTGACGCGAGGGCGCGCTTCTTTTGTCAAGACCGACTGTCGGGTGCC 3184
Qy 3895 CTGAATGAATGACGAGGACGCGAGCGGTATCGTGGTGGCCACGACGGCGGTTCCCT 3954
Db 3185 CTGAATGAATGACGAGGACGCGAGCGGTATCGTGGTGGCCACGACGGCGGTTCCCT 3244
Qy 3955 TCGCAGCTGTGCTCGAGCTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGCGAA 4014
Db 3245 TCGCAGCTGTGCTCGAGCTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGCGAA 3304

Qy 4015 GTCCCGGGCAGGATCTCCTGTTCATCTCACCTTGTCTCTCCGAGAAAGTATCCATCATG 4074
Db 3305 GTCCCGGGCAGGATCTCCTGTTCATCTCACCTTGTCTCTCCGAGAAAGTATCCATCATG 3364
Qy 4075 GCTGATCAATGGCGGCTGTCATACGCTTGTATCCGGCTACCTGCCCATTCGACACCAAA 4134
Db 3365 GCTGATCAATGGCGGCTGTCATACGCTTGTATCCGGCTACCTGCCCATTCGACACCAAA 3424
Qy 4135 GCGAAACATCGCATCGAGCGAGCAGTACTCGGANTGAAAGCCGGTCTTGTGATCAGGAT 4194
Db 3425 GCGAAACATCGCATCGAGCGAGCAGTACTCGGANTGAAAGCCGGTCTTGTGATCAGGAT 3484
Qy 4195 GATCTGACCAAGACATCAGGGGCTCGCGCCAGCCGAACTGTTGCGCAGCTCAAGGCG 4254
Db 3485 GATCTGACCAAGACATCAGGGGCTCGCGCCAGCCGAACTGTTGCGCAGCTCAAGGCG 3544
Qy 4255 CGCATGCCCGACCGCAGGATCTCGTGTGACCCATCGCGATGCTCTCTTGGCCGAATATC 4314
Db 3545 CGCATGCCCGACCGCAGGATCTCGTGTGACCCATCGCGATGCTCTCTTGGCCGAATATC 3604
Qy 4315 ATGGTGAATAATGGCCGCTTTTCTGGAATTCATCGACTGTGGCCGCTGGGTGTGGCGAC 4374
Db 3605 ATGGTGAATAATGGCCGCTTTTCTGGAATTCATCGACTGTGGCCGCTGGGTGTGGCGAC 3664
Qy 4375 CGCTATCAGGACATAGCGTTGGCTACCCGCTGATATTGCTGAAGAGCTTGGCGCGAATGG 4434
Db 3665 CGCTATCAGGACATAGCGTTGGCTACCCGCTGATATTGCTGAAGAGCTTGGCGCGAATGG 3724
Qy 4435 GCTGACCGCTTCTCTGCTGCTTTTACGGTATCGCGCTCCCGATTCGACGCGATCGCCTTC 4494
Db 3725 GCTGACCGCTTCTCTGCTGCTTTTACGGTATCGCGCTCCCGATTCGACGCGATCGCCTTC 3784
Qy 4495 TATCGCCTTCTTGACGAGTCTTCTTGACGCGGACTCTGGGGTTCGATAAATAAAGATT 4554
Db 3785 TATCGCCTTCTTGACGAGTCTTCTTGACGCGGACTCTGGGGTTCGATAAATAAAGATT 3844
Qy 4555 TTATTTAGTCTCCAGAAAAAGGGGGAATGAAGACCCACCTGTAGTGTGGCAAGCTA 4614
Db 3845 TTATTTAGTCTCCAGAAAAAGGGGGAATGAAGACCCACCTGTAGTGTGGCAAGCTA 3904
Qy 4615 GCTTAAGTAACGCGCATTTTTCGAAGGCATGGAAAAATACATAACTGAGAATAGAGAAGTTC 4674
Db 3905 GCTTAAGTAACGCGCATTTTTCGAAGGCATGGAAAAATACATAACTGAGAATAGAGAAGTTC 3964
Qy 4675 AGATCAAGGTCAGGAACAGATGGAAACAGCTGAATATGGGCCAAACACAGATATCTGTGTA 4734
Db 3965 AGATCAAGGTCAGGAACAGATGGAAACAGCTGAATATGGGCCAAACACAGATATCTGTGTA 4024
Qy 4735 AGCAGTTCTCCCGCTCAGGGCCAGAACAGATGGAACAGCTGAATATGGGCCAAAC 4794
Db 4025 AGCAGTTCTCCCGCTCAGGGCCAGAACAGATGGAACAGCTGAATATGGGCCAAAC 4084
Qy 4795 AGGATATCTGTGTAGCAGTCTCTGCCCGGCTCAGGGCCAGAACAGATGGTCCCCAG 4854
Db 4085 AGGATATCTGTGTAGCAGTCTCTGCCCGGCTCAGGGCCAGAACAGATGGTCCCCAG 4144
Qy 4855 ATGCGGTCCAGCCCTCAGCAGTCTTAGAGAAACATCAGATGTTTCCAGGGTGGCCCAAG 4914
Db 4145 ATGCGGTCCAGCCCTCAGCAGTCTTAGAGAAACATCAGATGTTTCCAGGGTGGCCCAAG 4204
Qy 4915 GACTGAAATGACCCCTGTGCTTATTTGAACTAACCAATCAGTTTCGCTTCTGT 4974
Db 4205 GACTGAAATGACCCCTGTGCTTATTTGAACTAACCAATCAGTTTCGCTTCTGT 4264
Qy 4975 TCGCGCGCTTCTGCTCCCGGAGCTCAATAAAGAGGCCACAAACCCCTCACTCGGGGGGCC 5034
Db 4265 TCGCGCGCTTCTGCTCCCGGAGCTCAATAAAGAGGCCACAAACCCCTCACTCGGGGGGCC 4324
Qy 5035 AGTCTCTCCGATTGATGAGTCGCGCGGGTACCGGTGTATCCAATAAACCCTCTTTCGAGTT 5094
Db 4325 AGTCTCTCCGATTGATGAGTCGCGCGGGTACCGGTGTATCCAATAAACCCTCTTTCGAGTT 4384

APPLICANT: University of Southern California
; TITLE OF INVENTION: Construction and Use of Genes Encoding Pathogenic Epitopes for Treatment of Autoimmune Disease
; FILE REFERENCE: 2013761-7030803001
; CURRENT APPLICATION NUMBER: US/10/359,397
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: 10/098,035
; PRIOR FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5856
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Sal/hind drag of pSVNA into Sal/Hind cut GI
US-10-359-397-1

Query Match 28.1%; Score 1455.2; DB 7; Length 5856;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 1488; Conservative 0; Mismatches 3; Indels 21; Gaps 1;

Qy 3687 ATCAAGATCTGATCAAGACAGAGATGAGGATCGTTTCGCAATGATGAACAAAGATGATT 3746
Db |||||
Qy 3747 GCAGCAGGTTCTCGGGCGCTTGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACA 3806
Db |||||
Qy 1904 GCAGCAGGTTCTCGGGCGCTTGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACA 1963
Db |||||
Qy 3807 GACAAATGGCTGCTGATGCGCGCTTCCGGCTGTCAGCGCAGGGCGCGGCTTCT 3866
Db |||||
Qy 1964 GACAAATGGCTGCTGATGCGCGCTTCCGGCTGTCAGCGCAGGGCGCGGCTTCT 2023
Db |||||
Qy 3867 TTTTGTCAAGACCGACCTGTCCGGTGCCTCGAATGAATGACGAGGAGGCGCGGCT 3926
Db |||||
Qy 2024 TTTTGTCAAGACCGACCTGTCCGGTGCCTCGAATGAATGACGAGGAGGCGCGGCT 2083
Db |||||
Qy 3927 ATCGTGGCTGCGCAGCAGCGGGCTTCTTGGCGAGCTGTGTCGACGTTGTCACCTGAAGC 3986
Db |||||
Qy 2084 ATCGTGGCTGCGCAGCAGCGGGCTTCTTGGCGAGCTGTGTCGACGTTGTCACCTGAAGC 2143
Db |||||
Qy 3987 GGGAAAGGACCTGGCTGCTATTGGGCGAAGTCCCGGGCAGGATCTCTGTCTCATCTCACCT 4046
Db |||||
Qy 2144 GGGAAAGGACCTGGCTGCTATTGGGCGAAGTCCCGGGCAGGATCTCTGTCTCATCTCACCT 2203
Db |||||
Qy 4047 TGCTCTCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGGCTGTCATACGCTTGA 4106
Db |||||
Qy 2204 TGCTCTCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGGCTGTCATACGCTTGA 2263
Db |||||
Qy 4107 TCCGGCTACCTGCCCATTCGACCAACCAAGCAACATCGCATCGAGCGAGCACGTAATCG 4166
Db |||||
Qy 2264 TCCGGCTACCTGCCCATTCGACCAACCAAGCAACATCGCATCGAGCGAGCACGTAATCG 2323
Db |||||
Qy 4167 GATGAAAGCGGCTCTTGTGATCAGGATGATCTGAGCAAGAGCATCAGGGGCTCGGCGC 4236
Db |||||
Qy 2324 GATGAAAGCGGCTCTTGTGATCAGGATGATCTGAGCAAGAGCATCAGGGGCTCGGCGC 2383
Db |||||
Qy 4227 AGCGGAATGTTTCGCGAGGCTCAAGGGCGCATGCCGAGCGGAGGATCTCGTCGTGAC 4286
Db |||||
Qy 2384 AGCGGAATGTTTCGCGAGGCTCAAGGGCGCATGCCGAGCGGAGGATCTCGTCGTGAC 2443
Db |||||
Qy 4287 CCATGGCGATGCTGCTTGGCGAATATCATGGTGGAAATGGCGCTTTTCTGGAATCAT 4346
Db |||||
Qy 2444 CCATGGCGATGCTGCTTGGCGAATATCATGGTGGAAATGGCGCTTTTCTGGAATCAT 2503
Db |||||
Qy 4347 CGACTGTGGCGGCTGGGTGGCGGCGCTATCAGGACATAGCGTTGGCTACCCGTA 4406
Db |||||
Qy 2504 CGACTGTGGCGGCTGGGTGGCGGCGCTATCAGGACATAGCGTTGGCTACCCGTA 2563
Db |||||
Qy 4407 TATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGC 4466
Db |||||
Qy 2564 TATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGC 2623
Db |||||

Qy 4467 CGCTCCGATTCGAGCGCATCGCTTCTATCGCTTCTTGACGAGTCTTCTCGAGCGG 4526
Db |||||
Qy 2624 CGCTCCGATTCGAGCGCATCGCTTCTATCGCTTCTTGACGAGTCTTCTCGAGCGG 2683
Db |||||
Qy 4527 ACTCTGGGGT-----TCGATAAAATAAAAGATTTTATTAGTCT 4565
Db |||||
Qy 2684 ACTCTGGGGTTCGTCGAGAGCTTGGGCCCATCGATAAAATAAAAGATTTTATTAGTCT 2743
Db |||||
Qy 4566 CAGAAAAAGGGGGAATGAAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAGTAAC 4625
Db |||||
Qy 2744 CAGAAAAAGGGGGAATGAAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAGTAAC 2803
Db |||||
Qy 4626 GCCATTTTCAAGGCATGGAAATAATACATAACTAGAGATAGAGAGTTTCAGATCAAGTTC 4685
Db |||||
Qy 2804 GCCATTTTCAAGGCATGGAAATAATACATAACTAGAGATAGAGAGTTTCAGATCAAGTTC 2863
Db |||||
Qy 4686 AGAAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAAGCAGTTCTCTG 4745
Db |||||
Qy 2864 AGAAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAAGCAGTTCTCTG 2923
Db |||||
Qy 4746 CCCCCTCAGGGCCAAAGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGT 4805
Db |||||
Qy 2924 CCCCCTCAGGGCCAAAGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGT 2983
Db |||||
Qy 4806 GGTAAAGCAGTTCTGCGCGCTCAGGGCCAAAGAACAGATGTTCCCCAGATGCGGTCCAG 4865
Db |||||
Qy 2984 GGTAAAGCAGTTCTGCGCGCTCAGGGCCAAAGAACAGATGTTCCCCAGATGCGGTCCAG 3043
Db |||||
Qy 4866 CCCTCAGCAGTTCTTAGAGAACCATCAGATGTTTCCAGGGTGCCCCAAAGACCTGAAATG 4925
Db |||||
Qy 3044 CCCTCAGCAGTTCTTAGAGAACCATCAGATGTTTCCAGGGTGCCCCAAAGACCTGAAATG 3103
Db |||||
Qy 4926 ACCTGTGCTTATTGTAACCAATCAGTTCGCTTCTCGCTTCTGTGCGCGGCTTC 4985
Db |||||
Qy 3104 ACCTGTGCTTATTGTAACCAATCAGTTCGCTTCTCGCTTCTGTGCGCGGCTTC 3163
Db |||||
Qy 4986 TGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACCTCGGGGCGCAGTCTCCGAT 5045
Db |||||
Qy 3164 TGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACCTCGGGGCGCAGTCTCCGAT 3223
Db |||||
Qy 5046 TGACTGAGTCCCGGATACCCGTGTATCAATAAAACCCCTTTCGAGTTGCATCCGACTT 5105
Db |||||
Qy 3224 TGACTGAGTCCCGGATACCCGTGTATCAATAAAACCCCTTTCGAGTTGCATCCGACTT 3283
Db |||||
Qy 5106 GTGGTCTCGCTTCTTGGGAGGCTCTCTCTGAGTGATGACTACCCGTCAGCGGGG 5165
Db |||||
Qy 3284 GTGGTCTCGCTTCTTGGGAGGCTCTCTCTGAGTGATGACTACCCGTCAGCGGGG 3343
Db |||||
Qy 5166 TCTTTTCATTGG 5177
Db |||||
Qy 3344 TCTTTTCATTGG 3355
Db |||||

RESULT 14
US-10-098-035-3
; Sequence 3, Application US/10098035
; Publication No. US20020141983A1
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; APPLICANT: Weiner, Leslie P.
; APPLICANT: McMillan, Minnie.
; TITLE OF INVENTION: CONSTRUCTION AND USE OF GENES ENCODING
; TITLE OF INVENTION: PATHOGENIC EPITOPES FOR TREATMENT OF AUTOIMMUNE DISEASE
; FILE REFERENCE: 13761-703-00 US
; CURRENT APPLICATION NUMBER: US/10/098,035
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: US/08/654,737
; PRIOR FILING DATE: 1996-05-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5865

[illegible]

Db 2639 ATTGAACAAGATGGATTGACGACAGGTTCTCCGGCGCTTGCGTGAGAGGCTATTTCGGC 2698
Qy 3790 TATGACTGGGCACACAGACAAATCGGCTGCTCTGATGCGCGCTGTTCCGGCTGTACGG 3849
Db 2699 TATGACTGGGCACACAGACAAATCGGCTGCTCTGATGCGCGCTGTTCCGGCTGTACGG 2758
Qy 3850 CAGGGCGCGCCGGTCTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAAATGAATGCGAG 3909
Db 2759 CAGGGCGCGCCGGTCTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAAATGAATGCGAG 2818
Qy 3910 GACGAGCAGCGCGCTATCGTGGCTGACGACGCGCGTTCCTGTCGCGAGCTGTGCTC 3969
Db 2819 GACGAGCAGCGCGCTATCGTGGCTGACGACGCGCGTTCCTGTCGCGAGCTGTGCTC 2878
Qy 3970 GACGTTGTCACTGAAGCGGGAAGGACTGGCTGTCTATTGGGCGAAGTCCCGGGCAGGAT 4029
Db 2879 GACGTTGTCACTGAAGCGGGAAGGACTGGCTGTCTATTGGGCGAAGTCCCGGGCAGGAT 2938
Qy 4030 TCTCTGTCACTCTACCTTGTCTCTCCCGAGAAAGTATCATCATGGCTGTATGCAATGCGG 4089
Db 2939 TCTCTGTCACTCTACCTTGTCTCTCCCGAGAAAGTATCATCATGGCTGTATGCAATGCGG 2998
Qy 4090 CGGCTGCATACGCTTGATCCGGCTACTGCGCATTCGACCAAGCGAAACATCGCATC 4149
Db 2999 CGGCTGCATACGCTTGATCCGGCTACTGCGCATTCGACCAAGCGAAACATCGCATC 3058
Qy 4150 GACGAGCAGCTPACTCGGATGGAGCGGCTTGTGCGATCAGGATGATCTGGACGAGAG 4209
Db 3059 GACGAGCAGCTPACTCGGATGGAGCGGCTTGTGCGATCAGGATGATCTGGACGAGAG 3118
Qy 4210 CATCAGGGGCTCGCCGACCGCAACTGTTGCGCAGGCTCAAGGCGCGCATGCCGACGGC 4269
Db 3119 CATCAGGGGCTCGCCGACCGCAACTGTTGCGCAGGCTCAAGGCGCGCATGCCGACGGC 3178
Qy 4270 GAGGATCTCGTGTGACCCATCGGCGATGCTGCTTGCCEAATATCATGGTGGAAAAATGGC 4329
Db 3179 GAGGATCTCGTGTGACCCATCGGCGATGCTGCTTGCCEAATATCATGGTGGAAAAATGGC 3238
Qy 4330 CGCTTTTCTGGATTATCATGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATA 4389
Db 3239 CGCTTTTCTGGATTATCATGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATA 3298
Qy 4390 GGGTTGGCTACCGGTGATATTGTGAAGAGCTTGGCGCGAATGGGCTGACCGCTTCTCCTC 4449
Db 3299 GGGTTGGCTACCGGTGATATTGTGAAGAGCTTGGCGCGAATGGGCTGACCGCTTCTCCTC 3358
Qy 4450 GTGCTTTACGGTATCGCGCTCCGATTCGACGCGCATCGGCTTCTATCGCCTTCTTGAC 4509
Db 3359 GTGCTTTACGGTATCGCGCTCCGATTCGACGCGCATCGGCTTCTATCGCCTTCTTGAC 3418
Qy 4510 GAGTTCTTCTGAGCGGAGCTCTGGGGTTGATATAAATAAAGATTTTATTAGTCTCCAG 4569
Db 3419 GAGTTCTTCTGAGCGGAGCTCTGGGGTTGATATAAATAAAGATTTTATTAGTCTCCAG 3478
Qy 4570 AAAAAGGGGGAAATGAAAGACCCACCTGTAGTTTGGCAAGCTAGCTTAAAGTAAACGCCA 4629
Db 3479 AAAAAGGGGGAAATGAAAGACCCACCTGTAGTTTGGCAAGCTAGCTTAAAGTAAACGCCA 3538
Qy 4630 TTTTTCGAAGGCATGGAATAAATACATACTAGAGAATAGAGAAGTTTCAGATCAAGGTCAAGGA 4689
Db 3539 TTTTTCGAAGGCATGGAATAAATACATACTAGAGAATAGAGAAGTTTCAGATCAAGGTCAAGGA 3598
Qy 4690 ACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTCTTGCCCC 4749
Db 3599 ACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTCTTGCCCC 3658
Qy 4750 GGCCTCAGGCGCAAGAACAGATGGACACGTGAATATGGGCCAAACAGGATATCTGTGGTA 4809
Db 3659 GGCCTCAGGCGCAAGAACAGATGGACACGTGAATATGGGCCAAACAGGATATCTGTGGTA 3718
Qy 4810 AGCAGTTCTTCCCGCGCTCAGGGCCAAAGAACAGATGGTCCCGCAGATCGGTCAGGCCCT 4869
Db 3719 AGCAGTTCTTCCCGCGCTCAGGGCCAAAGAACAGATGGTCCCGCAGATCGGTCAGGCCCT 3778

Qy 4870 CAGCAGTTTCTAGAGAAACCATCAGATGTTTCCAGGGTGCCCCAAGGACCTGAAATGACCC 4929
Db 3779 CAGCAGTTTCTAGAGAAACCATCAGATGTTTCCAGGGTGCCCCAAGGACCTGAAATGACCC 3838
Qy 4930 TGTGCTTTATTGAACTAAACCAATCAGTTGCTTCTCGCTTCTGTTGCGGCGCTTCTGCT 4989
Db 3839 TGTGCTTTATTGAACTAAACCAATCAGTTGCTTCTCGCTTCTGTTGCGGCGCTTCTGCT 3898
Qy 4990 CCCCAGGCTCAATAAAAGAGCCCAAAACCCCTCACTCGGGGCGCAGTCTCCGATTGAC 5049
Db 3899 CCCCAGGCTCAATAAAAGAGCCCAAAACCCCTCACTCGGGGCGCAGTCTCCGATTGAC 3958
Qy 5050 TGAGTCGCCCGGGTACCCGTGTATCCAAATAAACCCCTCTTGCAATTGCATCCGACTTGTGG 5109
Db 3959 TGAGTCGCCCGGGTACCCGTGTATCCAAATAAACCCCTCTTGCAATTGCATCCGACTTGTGG 4018
Qy 5110 TCTGCTGTTCTCTTGGGAGGGTCTCTCTGAGTGAATTGACTACCCGTCAGCGGGGCTT 5169
Db 4019 TCTGCTGTTCTCTTGGGAGGGTCTCTCTGAGTGAATTGACTACCCGTCAGCGGGGCTT 4078
Qy 5170 TCATTGG 5177
Db 4079 TCATTGG 4086

Search completed: February 27, 2006, 10:26:20
Job time : 3831 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OW nucleic - nucleic search, using sw model

Run on: February 27, 2006, 08:43:40 ; Search time 1260 Seconds
(without alignments)
8761.320 Million cell updates/sec

Title: US-10-618-570-1

Perfect score: 5177

Sequence: 1 ttgaagaccaccaccgta.....agcgggggtttttcattgg 5177

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7209121 seqs, 1066183437 residues

Total number of hits satisfying chosen parameters: 14418242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA New:

- 1: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq*
- 2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq*
- 3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq*
- 4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq*
- 5: /cgn2_6/ptodata/1/pubpna/US03_NEW_PUB.seq*
- 6: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq*
- 7: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq*
- 8: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq*
- 9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq*
- 10: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2*
- 11: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3*
- 12: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq4*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	1170.4	22.6	3974	12 US-11-205-031-10	Sequence 10, Appl
C 2	1170.4	22.6	3974	12 US-11-233-119-16	Sequence 16, Appl
3	1170.4	22.6	12445	7 US-10-948-344-2	Sequence 2, Appl
C 4	1167.8	22.6	2290	12 US-11-053-187-16	Sequence 16, Appl
5	1166.2	22.5	2403	7 US-10-523-682-1	Sequence 1, Appl
C 6	1154.4	22.3	8136	12 US-11-082-154A-31	Sequence 31, Appl
C 7	1154.4	22.3	17384	12 US-11-082-154A-118	Sequence 18, Appl
8	1136.4	22.0	4323	8 US-10-525-710-66	Sequence 66, Appl
9	1136.4	22.0	4323	8 US-10-525-674-58	Sequence 58, Appl
10	1136.4	22.0	5091	8 US-10-525-710-65	Sequence 65, Appl
11	1136.4	22.0	5091	8 US-10-525-674-57	Sequence 57, Appl
12	1136.4	22.0	5860	8 US-10-525-710-69	Sequence 69, Appl
13	1136.4	22.0	5860	8 US-10-525-710-74	Sequence 74, Appl
14	1136.4	22.0	5860	8 US-10-525-674-61	Sequence 61, Appl
15	1136.4	22.0	5860	8 US-10-525-674-66	Sequence 66, Appl
16	1136.4	22.0	6472	8 US-10-525-674-69	Sequence 69, Appl
17	1136.4	22.0	6591	8 US-10-525-710-79	Sequence 79, Appl
18	1120.8	21.6	3906	12 US-11-053-187-17	Sequence 17, Appl
19	1120.8	21.6	5856	7 US-10-933-746-34	Sequence 34, Appl
20	1120.8	21.6	7448	12 US-11-071-651-15	Sequence 15, Appl

21	1089.6	21.0	6233	12 US-11-193-750-10	Sequence 10, Appl
22	1088	21.0	5225	12 US-11-231-725-1	Sequence 1, Appl
C 23	971	18.8	3534	12 US-11-186-282-22	Sequence 22, Appl
C 24	970.2	18.7	2700	12 US-11-186-282-33	Sequence 33, Appl
C 25	970.2	18.7	2716	12 US-11-186-282-27	Sequence 27, Appl
C 26	970.2	18.7	2716	12 US-11-186-282-28	Sequence 28, Appl
C 27	970.2	18.7	2716	12 US-11-186-282-29	Sequence 29, Appl
C 28	970.2	18.7	2716	12 US-11-186-282-31	Sequence 31, Appl
C 29	970.2	18.7	2721	12 US-11-186-282-26	Sequence 26, Appl
C 30	970.2	18.7	2721	12 US-11-186-282-34	Sequence 34, Appl
C 31	970.2	18.7	2722	12 US-11-186-282-23	Sequence 23, Appl
C 32	970.2	18.7	2725	12 US-11-186-282-24	Sequence 24, Appl
C 33	970.2	18.7	2725	12 US-11-186-282-25	Sequence 25, Appl
C 34	970.2	18.7	2725	12 US-11-186-282-32	Sequence 32, Appl
C 35	970.2	18.7	2739	12 US-11-186-282-30	Sequence 30, Appl
C 36	957.2	18.5	5181	12 US-11-094-484-3	Sequence 3, Appl
C 37	863.6	16.7	8074	7 US-10-493-909-14	Sequence 14, Appl
38	856.2	16.5	6602	7 US-10-493-909-100	Sequence 100, App
39	851.6	16.4	4465	7 US-10-493-909-13	Sequence 13, Appl
40	850.4	16.4	7127	12 US-11-192-219-46	Sequence 46, Appl
41	850.4	16.4	7201	12 US-11-159-919-16	Sequence 16, Appl
42	850.4	16.4	7496	12 US-11-159-919-15	Sequence 15, Appl
43	849.8	16.4	12604	12 US-11-137-315A-44	Sequence 44, Appl
44	849.8	16.4	12604	12 US-11-180-074-39	Sequence 39, Appl
45	849.8	16.4	14561	12 US-11-137-315A-47	Sequence 47, Appl

ALIGNMENTS

RESULT 1

US-11-205-031-10/c
; Sequence 10, Application US/11205031
; Publication No. US20050287588A1
; GENERAL INFORMATION:
; APPLICANT: Ji et al.
; TITLE OF INVENTION: Breast Cancer Specific Gene 1
; FILE REFERENCE: PF342D1
; CURRENT APPLICATION NUMBER: US/11/205,031
; CURRENT FILING DATE: 2005-08-17
; PRIOR APPLICATION NUMBER: US 09/017,715
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: US 60/037,080
; PRIOR FILING DATE: 1997-02-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 3974
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-205-031-10

Query Match	22.6%	Score 1170.4;	DB 12;	Length 3974;
Best Local Similarity	99.9%	Pred. No. 8.7e+268;		
Matches 1171;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	3371	AAGTTCACGCTGCGCGCAAGCAGCTCAGGCGCAGGAGGCTGCTAAAGGAAGGGAACAGCT	3430	
Db	3868	AGCTTCACGCTGCGCGCAAGCAGCTCAGGCGCAGGAGGCTGCTAAAGGAAGGGAACAGCT	3809	
Qy	3431	AGAAAGCCAGTCCGCGCAAGAAACGGTGTGACCCCGATGAATGTGAGTCTGAGCTTATCT	3490	
Db	3808	AGAAAGCCAGTCCGCGCAAGAAACGGTGTGACCCCGATGAATGTGAGTCTGAGCTTATCT	3749	
Qy	3491	GGACAGGGAAGAAACGCGCAAGGCGCAAGAAAGAGAGGAGTGTGAGTGGGCTTACATGGC	3550	
Db	3748	GGACAGGGAAGAAACGCGCAAGGCGCAAGGAGGAGTGTGAGTGGGCTTACATGGC	3689	
Qy	3551	GATAGCTAGACTGGGCGGTTTATGGACAGCAAGCGGAATTCAGCTGGGGCGC	3610	
Db	3688	GATAGCTAGACTGGGCGGTTTATGGACAGCAAGCGGAATTCAGCTGGGGCGC	3629	
Qy	3611	CCTCTGTAAGTTGGGAAGCCCTGCAAGTAAACTGGATGGCTTCTTCTCGCCCAAGGA	3670	

```
Db CCTCTGGTAAAGTTGGGAAGCCCTGCAAGTAAGTGAATGCTTCTTTCGCGCAAGGA 3569
Qy TCTGATGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGA 3730
Db TCTGATGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGA 3509
Qy TTGACACAGATGGAATGACACAGGTTCTCCGCGCGCTTGGGTGGAGAGGCTATTCGGCT 3790
Db TTGAACAAGATGGAATGACACAGGTTCTCCGCGCGCTTGGGTGGAGAGGCTATTCGGCT 3449
Qy ATGACTGGGCAACACAGCAATCGGCTGCTCTGATGCGCGCGCTTTCGGGCTGTCAGGCG 3850
Db ATGACTGGGCAACACAGCAATCGGCTGCTCTGATGCGCGCGCTTTCGGGCTGTCAGGCG 3389
Qy AGGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTTCGGTGCCTGAATGAATGCACTGCAGG 3910
Db AGGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTTCGGTGCCTGAATGAATGCACTGCAGG 3329
Qy ACGAGGACGCGGCTATCGTGGCTGCGCACGACGCGGCTTCTTTCGGCAGCTGTGCTCG 3970
Db ACGAGGACGCGGCTATCGTGGCTGCGCACGACGCGGCTTCTTTCGGCAGCTGTGCTCG 3269
Qy ACGTTGTCACTGAAGCGGGAAGGACTGGCTGCTATTTGGGCAAGTGCCTGGGCGAGGATC 4030
Db ACGTTGTCACTGAAGCGGGAAGGACTGGCTGCTATTTGGGCGAAGTGCCTGGGCGAGGATC 3209
Qy TCTGTCTATCACTTCTGCTGCGGAGAAAGTATTCATCATGCTGATGCAATGCGGC 4090
Db TCTGTCTATCACTTCTGCTGCGGAGAAAGTATTCATCATGCTGATGCAATGCGGC 3149
Qy GGCTGCATACGTTGATCGGCTACCTGCCATTTGACACCAAGCGAAATCGCATCG 4150
Db GGCTGCATACGTTGATCGGCTACCTGCCATTTGACACCAAGCGAAATCGCATCG 3089
Qy AGCGAGCAGTACTCGGATGGAAGCGGCTTGTGTCGATCAGGATGATCTGACCAAGAGC 4210
Db AGCGAGCAGTACTCGGATGGAAGCGGCTTGTGTCGATCAGGATGATCTGACCAAGAGC 3029
Qy ATCAGGGGCTCGCGCAGCCGAACTGTTTCGCAAGCTCAAGCGCGCATGCCGACGCG 4270
Db ATCAGGGGCTCGCGCAGCCGAACTGTTTCGCAAGCTCAAGCGCGCATGCCGACGCG 2969
Qy AGGATCTGTCGTGACCCATGCGCATGCTGCTTCCGCAATATCATGCTGAAATGCGC 4330
Db AGGATCTGTCGTGACCCATGCGCATGCTGCTTCCGCAATATCATGCTGAAATGCGC 2909
Qy GCTTTTCTGGATTTCATCGACTGTGGCGGCTGGGTGTGGCGACCGCTATCAGGACATAG 4390
Db GCTTTTCTGGATTTCATCGACTGTGGCGGCTGGGTGTGGCGACCGCTATCAGGACATAG 2849
Qy CGTTGGCTACCGTGATATGCTGAAGAGCTTGGCGGCAATGGGCTGACCGCTTCCTCG 4450
Db CGTTGGCTACCGTGATATGCTGAAGAGCTTGGCGGCAATGGGCTGACCGCTTCCTCG 2789
Qy TCGTTTACGTTATCGCGCTCCGATTCGACGCGATCGCTTCTATCGCTTCTTGACG 4510
Db TCGTTTACGTTATCGCGCTCCGATTCGACGCGATCGCTTCTATCGCTTCTTGACG 2729
Qy AGTTCTTCTGAGCGGACTCTGGGTTTCGATA 4542
Db AGTTCTTCTGAGCGGACTCTGGGTTTCGATA 2697
```

RESULT 2

```
US-11-233-119-16/c
; Sequence 16, Application US/11233119
; Publication No. US20060025331A1
; GENERAL INFORMATION:
; APPLICANT: HU, JING-SHAN et al,
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; FILE REFERENCE: PF112P4D1
; CURRENT APPLICATION NUMBER: US/11/233,119
```

```
; CURRENT FILING DATE: 2005-09-23
; PRIOR APPLICATION NUMBER: 09/107,997
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: 09/042,105
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 09/999,811
; PRIOR FILING DATE: 1997-12-24
; PRIOR APPLICATION NUMBER: 08/465,968
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/207,550
; PRIOR FILING DATE: 1994-03-08
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 3974
; TYPE: DNA
; ORGANISM: Expression vector pHEA4-5
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(3974)
; OTHER INFORMATION: Expression vector pHEA4-5
; US-11-233-119-16
```

```
Query Match 22.6%; Score 1170.4; DB 12; Length 3974;
Best Local Similarity 99.9%; Pred. No. 8.7e-268;
Matches 1171; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 3371 AAGTTTCACGTCGCGCAAGCACTCAGGGCGCAAGGGCTGCTAAAGGAAGCGGAACAGCT 3430
Db 3868 AAGTTTCACGTCGCGCAAGCACTCAGGGCGCAAGGGCTGCTAAAGGAAGCGGAACAGCT 3809
Qy 3431 AGAAGCCAGTCCGCAAGACGGTGTGACCCCGGATGAATGTGAGTCTTCTGAGTCTGCT 3490
Db 3808 AGAAGCCAGTCCGCAAGACGGTGTGACCCCGGATGAATGTGAGTCTTCTGAGTCTGCT 3749
Qy 3491 GGACACGGGAAACCGCAAGCGCAAGAAAGAGAGAGAGTAGCTTGCAGTGGGCTTACATGGC 3550
Db 3748 GGACACGGGAAACCGCAAGCGCAAGAAAGAGAGAGTAGCTTGCAGTGGGCTTACATGGC 3689
Qy 3551 GATAGCTAGACTGGCGGTTTTATGGAACAGCAAGCAAGCGGAATTCGAGCTGGGCGC 3610
Db 3688 GATAGCTAGACTGGCGGTTTTATGGAACAGCAAGCAAGCGGAATTCGAGCTGGGCGC 3629
Qy 3611 CCTCTGTAGGTTGGGAAGCCCTGCAAGTAACTGGATGGCTTCTTCCGCGCAAGGA 3670
Db 3628 CCTCTGTAGGTTGGGAAGCCCTGCAAGTAACTGGATGGCTTCTTCCGCGCAAGGA 3569
Qy 3671 TCTGATGCGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGA 3730
Db 3568 TCTGATGCGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGA 3509
Qy 3731 TTGAACAAGATGGAATTCGACGAGGTTCTCCGCGCGCTTGGGTGGAGAGGCTATTCGGCT 3790
Db 3508 TTGAACAAGATGGAATTCGACGAGGTTCTCCGCGCGCTTGGGTGGAGAGGCTATTCGGCT 3449
Qy 3791 ATGACTGGGCAACACAGCAATCGGCTGCTGATGCGCGCTTCCGCTGTCAGGCG 3850
Db 3448 ATGACTGGGCAACACAGCAATCGGCTGCTGATGCGCGCTTCCGCTGTCAGGCG 3389
Qy 3851 AGGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGCTCCGGTGCCTGAACTGAATGCAAG 3910
Db 3388 AGGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGCTCCGGTGCCTGAACTGAATGCAAG 3329
Qy 3911 ACGAGGACGCGGCTATCGTGGCTGCGCACAGCGGGGCTTCTTTCGGCAGCTGTGCTCG 3970
Db 3328 ACGAGGACGCGGCTATCGTGGCTGCGCACAGCGGGGCTTCTTTCGGCAGCTGTGCTCG 3269
Qy 3971 ACGTTGTCACTGAAGCGGGAAGGACTGGCTGCTATTTGGGCGAAGTGCCTGGGCGAGGATC 4030
Db 3268 ACGTTGTCACTGAAGCGGGAAGGACTGGCTGCTATTTGGGCGAAGTGCCTGGGCGAGGATC 3209
Qy 4031 TCCTGTCTATCACTTCTGCTGCGGAGAAAGTATTCATCATGCTGATGCAATGCGGC 4090
```

3208	TCCTGTCAATCTCACCTTGCTCTCGCGGAGAAAGTATCCATCATGTGGCTGTGATGCAATGCGGC	3149
4091	GGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCAACGAAGGAAACATCGCATCG	4150
3148	GGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCAACGAAGGAAACATCGCATCG	3089
4151	AGCGAGCACGTACTCGGATCGGAGCGGGTCTTGTGATCAGGATGATCTGGACGAAGGC	4210
3088	AGCGAGCACGTACTCGGATCGGAGCGGGTCTTGTGATCAGGATGATCTGGACGAAGGC	3029
4211	ATCAGGGGCTCGGCGCCAGCGGAACTGTTTGCACAGGCTCAAGGGCGCGCATGCCCGAAGCGG	4270
3028	ATCAGGGGCTCGGCGCCAGCGGAACTGTTTGCACAGGCTCAAGGGCGCGCATGCCCGAAGCGG	2969
4271	AGGATCTCGTGTGACCCATGGCGATGCCCTGTTCGGGAATATCATGGTGGAAAAATGGCC	4330
2968	AGGATCTCGTGTGACCCATGGCGATGCCCTGTTCGGGAATATCATGGTGGAAAAATGGCC	2909
4331	GCTTTTCTGGAATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAG	4390
2908	GCTTTTCTGGAATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAG	2849
4391	CGTTGGCTACCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCTCG	4450
2848	CGTTGGCTACCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCTCG	2789
4451	TGCTTTTACGGTATCGCCGCTCCGATTCGACAGCGCATCGCTTCTATCGGCTTCTTTGACG	4510
2788	TGCTTTTACGGTATCGCCGCTCCGATTCGACAGCGCATCGCTTCTATCGGCTTCTTTGACG	2729
4511	AGTTCTTCTGAGCGGGAATCTGGGGTTCGATA	4542
2728	AGTTCTTCTGAGCGGGAATCTGGGGTTCGAAA	2697

RESULT 3

```

US-10-948-344-2
; Sequence 2, Application US/10948344
; Publication No. US20060013805A1
; GENERAL INFORMATION:
; APPLICANT: Hebbel, R.P.
; APPLICANT: Lin, Y.
; APPLICANT: Lollar, J.S.
; TITLE OF INVENTION: Transgenic circulating endothelial cells
; FILE REFERENCE: 600.449US1
; CURRENT APPLICATION NUMBER: US/10/948,344
; CURRENT FILING DATE: 2004-09-23
; PRIOR APPLICATION NUMBER: US/03/865,022
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: PCT/US99/28033
; PRIOR FILING DATE: 1999-11-24
; PRIOR APPLICATION NUMBER: US 60/109,687
; PRIOR FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 12445
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: The DNA sequence of HSQREneo.
US-10-948-344-2

```

	Query Match	22.6%	Score 1170.4	DB 7	Length 12445
	Best Local Similarity	99.9%	Pred. No. 1.5e-267		
	Matches 1171	Conservative 0	Mismatches 1	Indels 0	Gaps 0
Qy	3371	AA GTTCA CCGCTG CCGCA AGCTC	AG GCGCG CAA GGGCTG CT	TA AAGGA ACGGAA CACGT	3430
Db	7480	AA GTTCA CCGCTG CCGCA AGCTC	AG GCGCG CAA GGGCTG CT	TA AAGGA ACGGAA CACGT	7539
Qy	3431	AG AAGCA CAGTCC GCGCA AAGCG	TG TGTGA CCCC GCGGATGA AT	TG TCACTA CTGGGCTATCT	3490

Db	7540	AGAAAGCCAGTCCGCAAGAAACGGTGTGTGACCCCGGATGAATGTACAGTACTGGGGCTATCT	7599
Qy	3491	GGACAAGGGAAAAACGCAAGCGCAAAAGAGAAAGCAGGTAGCTTGCAGTGGGCTTACATGGC	3550
Db	7600	GGACNAGGGNAAACGCAAGCGCAAAAGNAAACAGGTAGCTTGCAGTGGGCTTACATGGC	7659
Qy	3551	GATAGCTAGACTCGGCGGTTTATGACAGCAAGCGAAACCGGAATTCAGCTTGGGGCGC	3610
Db	7660	GATAGCTAGACTCGGCGGTTTATGACAGCAAGCGAAACCGGAATTCAGCTTGGGGCGC	7719
Qy	3611	CCCTCTGGTAGGTTGGNAGCCCTGCAAGTAAACTGGATGGCTTCTTGGCCGCCAAGCA	3670
Db	7720	CCCTCTGGTAGGTTGGNAGCCCTGCAAGTAAACTGGATGGCTTCTTGGCCGCCAAGCA	7779
Qy	3671	TCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGGATGAGGATGTTTCGCATGA	3730
Db	7780	TCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGGATGAGGATGTTTCGCATGA	7839
Qy	3731	TTGAAACAAGATGGAATTGCACGCAAGTTCTCCGGCCGCTTGGGTFGAGAGGCTATTCCGGCT	3790
Db	7840	TTGAAACAAGATGGAATTGCACGCAAGTTCTCCGGCCGCTTGGGTFGAGAGGCTATTCCGGCT	7899
Qy	3791	ATGACTGGGCACAACAGACAAATCGGCTGCTCTGATGCGCCGCTGTTCCGGCTGTCAGCC	3850
Db	7900	ATGACTGGGCACAACAGACAAATCGGCTGCTCTGATGCGCCGCTGTTCCGGCTGTCAGCC	7959
Qy	3851	AGGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCGGTGGCCCTGAAATGAACCTCAGCG	3910
Db	7960	AGGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCGGTGGCCCTGAAATGAACCTCAGCG	8019
Qy	3911	ACGAGGACGCGCGGCTATCGTGGCTGGCCACGACGGCGGTTCTTCCGCGAGCTGTGCTCG	3970
Db	8020	ACGAGGACGCGCGGCTATCGTGGCTGGCCACGACGGCGGTTCTTCCGCGAGCTGTGCTCG	8079
Qy	3971	ACGTTGTCACTGAAGCGGGNAGGGACTGGCTGCTATTTGGCGNAGTGCCTGGGCGCAGATC	4030
Db	8080	ACGTTGTCACTGAAGCGGGNAGGGACTGGCTGCTATTTGGCGNAGTGCCTGGGCGCAGATC	8139
Qy	4031	TCCTGTCTCATCTCACTTGTCTCTGCGGAGAAAGTATCCATCATGGCTGATGCAATCGCGC	4090
Db	8140	TCCTGTCTCATCTCACTTGTCTCTGCGGAGAAAGTATCCATCATGGCTGATGCAATCGCGC	8199
Qy	4091	GGCTGCATACGCTTGTATCGGGCTACCTGCCCATTCGACCAACCAAGCGAAACATCGCATCG	4150
Db	8200	GGCTGCATACGCTTGTATCGGGCTACCTGCCCATTCGACCAACCAAGCGAAACATCGCATCG	8259
Qy	4151	AGCGAGCATGTACTCGGATGGAAAGCCGGTCTTGTTCGATCAGAGTATCTGGACGAAGAGC	4210
Db	8260	AGCGAGCATGTACTCGGATGGAAAGCCGGTCTTGTTCGATCAGAGTATCTGGACGAAGAGC	8319
Qy	4211	ATCAGGGGCTCGCGCAGCGCAACTGTTGCGCAGGCTCAAGGCGCGCATGCCACGGCG	4270
Db	8320	ATCAGGGGCTCGCGCAGCGCAACTGTTGCGCAGGCTCAAGGCGCGCATGCCACGGCG	8379
Qy	4271	AGGATCTGTCGTGACCCATGGCGATCGCTTGTCCGGAATATCATGGTGGAAAAATGGCC	4330
Db	8380	AGGATCTGTCGTGACCCATGGCGATCGCTTGTCCGGAATATCATGGTGGAAAAATGGCC	8439
Qy	4331	GCCTTTCTGGAATTCATGCTGTGGCGGGTGTGGGAGCCGCTATCAGGACATAG	4390
Db	8440	GCCTTTCTGGAATTCATGCTGTGGCGGGTGTGGGAGCCGCTATCAGGACATAG	8499
Qy	4391	CGTTGGCTACCGGTGATTTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCTCG	4450
Db	8500	CGTTGGCTACCGGTGATTTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCTCG	8559
Qy	4451	TGCTTTACGATTCGCGGCTCCCGATTCGACGGCATCGCCTTCTATCGCCTTCTTTGAGC	4510
Db	8560	TGCTTTACGATTCGCGGCTCCCGATTCGACGGCATCGCCTTCTATCGCCTTCTTTGAGC	8619
Qy	4511	AGTTCTTCTGAGCGGACTCTGGGGTTTCGATA	4542
Db	8620	AGTTCTTCTGAGCGGACTCTGGGGTTTCGAAA	8651

RESULT 4

US-11-053-187-16/c
; Sequence 16, Application US/11053187
; Publication No. US20050282184A1
; GENERAL INFORMATION:
; APPLICANT: INVITROGEN CORPORATION
; APPLICANT: CHESNUT, Jonathan D.
; APPLICANT: SHUMAN, Stewart
; APPLICANT: MADDEN, Knut R.
; APPLICANT: HEYMAN, John A.
; APPLICANT: BENNETT, Robert P.
; TITLE OF INVENTION: METHODS AND REAGENTS FOR MOLECULAR CLONING
; FILE REFERENCE: INVIT1300-1
; CURRENT APPLICATION NUMBER: US/11/053,187
; CURRENT FILING DATE: 2005-02-07
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US/09/935,280
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,563
; PRIOR FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 2290
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: puni/V5-His version A vector
US-11-053-187-16

Query Match 22.6%; Score 1167.8; DB 12; Length 2290;

Best Local Similarity 99.8%; Pred. No. 2.8e-267;

Matches 1169; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	3372	AGCTTCACGCTGCCGCAAGCACTCAGGCGCAAGGGCTGCTAAAGGAAGCGGAACAGTA	3431
Db	2290	AGCTTCACGCTGCCGCAAGCACTCAGGCGCAAGGGCTGCTAAAGGAAGCGGAACAGTA	2231
Qy	3432	GAAGGCCAGTCCGCAAGAACGGTGTGACCCCGGATGAATCTCAGCTACTGGGCTATCTG	3491
Db	2230	GAAGGCCAGTCCGCAAGAACGGTGTGACCCCGGATGAATCTCAGCTACTGGGCTATCTG	2171
Qy	3492	GACAAGGGAACCGCAAGCGCAAGAGAACAGCAGGTAGCTGCAAGTGGGCTTACATGGCG	3551
Db	2170	GACAAGGGAACCGCAAGCGCAAGAGAACAGCAGGTAGCTGCAAGTGGGCTTACATGGCG	2111
Qy	3552	ATAGCTAGCTGGGCGGTTTATGACACAGCAAGCGAACCGGAATGGCCAGCTGGGGCGCC	3611
Db	2110	ATAGCTAGCTGGGCGGTTTATGACACAGCAAGCGAACCGGAATGGCCAGCTGGGGCGCC	2051
Qy	3612	CTCTGGTAAGGTTGGGAAGCCCTGCAAGAGTAACTGGATGGCTTTCTTGGCGCAAGGAT	3671
Db	2050	CTCTGGTAAGGTTGGGAAGCCCTGCAAGAGTAACTGGATGGCTTTCTTGGCGCAAGGAT	1991
Qy	3672	CTGATGGCGCAGGGGATCAAGATCTGATCAAGACAGGATGAGGATCGTTTCCATGAT	3731
Db	1990	CTGATGGCGCAGGGGATCAAGATCTGATCAAGACAGGATGAGGATCGTTTCCATGAT	1931
Qy	3732	TGAACAGATGGATTGCAAGCAGGTTCTCCGGCGCTTGGGTGAGAGGCTATTTCGGCTA	3791
Db	1930	TGAACAGATGGATTGCAAGCAGGTTCTCCGGCGCTTGGGTGAGAGGCTATTTCGGCTA	1871
Qy	3792	TGACTGGGCACACAGACAATCGGCTGCTGTGATGCGCGCGTGTTCGGGCTGTCCAGGCA	3851
Db	1870	TGACTGGGCACACAGACAATCGGCTGCTGTGATGCGCGCGTGTTCGGGCTGTCCAGGCA	1811
Qy	3852	GGGGCGCCCGGTTCTTTTGTCAAGACCGACTGTCCGGTCCCTGATGAATGCAATGCAAG	3911
Db	1810	GGGGCGCCCGGTTCTTTTGTCAAGACCGACTGTCCGGTCCCTGATGAATGCAATGCAAG	1751
Qy	3912	CGAGGCGCGCGGCTATCGTGCTGGCTGGCCACGACGCGGGCTTCCTTGGCGAGCTGTGCTCGA	3971

Db	1750	CGAGGCGCGCGCTATCGTGGCTTGCCACGACGGGGTTCCTTGGCGAGCTGTGCTCGA	1691
Qy	3972	CGTTGTCACTGAAGCGGGAAGGAGCTGGCTGCTATTGGGCGAAGTGGCGGCGAGGATCT	4031
Db	1690	CGTTGTCACTGAAGCGGGAAGGAGCTGGCTGCTATTGGGCGAAGTGGCGGCGAGGATCT	1631
Qy	4032	CTGTGATCTCACCTTCTCTCCGAGAAAGTATCCATCATGCTCATGCAATGCGGCG	4091
Db	1630	CTGTGATCTCACCTTCTCTCCGAGAAAGTATCCATCATGCTCATGCAATGCGGCG	1571
Qy	4092	GCTGCATATCCTTGTATCCGGCTACCTGCGCCATTTCGACCAAGCGAAGCAATCGCATCGA	4151
Db	1570	GCTGCATATCCTTGTATCCGGCTACCTGCGCCATTTCGACCAAGCGAAGCAATCGCATCGA	1511
Qy	4152	GCAGACACGTACTCGGATGGAAGCCGGTCTTTGCGATCAGGATGATCTGGACGAAGACA	4211
Db	1510	GCAGACACGTACTCGGATGGAAGCCGGTCTTTGCGATCAGGATGATCTGGACGAAGACA	1451
Qy	4212	TCAGGGGCTCGCCGACCGCAACTGTTGCGCAGGCTCAAGGCGGCGATGCCGACGGCGA	4271
Db	1450	TCAGGGGCTCGCCGACCGCAACTGTTGCGCAGGCTCAAGGCGGCGATGCCGACGGCGA	1391
Qy	4272	GGATCTCGTGTGACCCATGCGCATGCTGCTTGCCTCCGAATATCATGTGGAAATGCGCG	4331
Db	1390	GGATCTCGTGTGACCCATGCGCATGCTGCTTGCCTCCGAATATCATGTGGAAATGCGCG	1331
Qy	4332	CTTTTCTGGATTCATGACATGCTGGCGCGCTGGGTGTGGCGGACCGCTATCAGGACATAGC	4391
Db	1330	CTTTTCTGGATTCATGACATGCTGGCGCGCTGGGTGTGGCGGACCGCTATCAGGACATAGC	1271
Qy	4392	GTGGGCTACCGGTGATATTCGTAAGAGCTTGGCGGGAATGGGCTGACCGCTTCTTCGT	4451
Db	1270	GTGGGCTACCGGTGATATTCGTAAGAGCTTGGCGGGAATGGGCTGACCGCTTCTTCGT	1211
Qy	4452	GCTTTACGGTATCGCGCTCCCGATTCGACGCGCATCGCCTTCTATCGCCTTCTTGACGA	4511
Db	1210	GCTTTACGGTATCGCGCTCCCGATTCGACGCGCATCGCCTTCTATCGCCTTCTTGACGA	1151
Qy	4512	GTTCTTCTGAGCGGAGCTCTGGGGTTCGATA	4542
Db	1150	GTTCTTCTGAGCGGAGCTCTGGGGTTCGATA	1120

RESULT 5

US-10-523-682-1
; Sequence 1, Application US/10523682
; Publication No. US20060014149A1
; GENERAL INFORMATION:
; APPLICANT: Schnorr, Kirk
; TITLE OF INVENTION: Methods for rolling circle amplification and signal trapping of
; TITLE OF INVENTION: Libraries
; FILE REFERENCE: 10292.204-US
; CURRENT APPLICATION NUMBER: US/10/523,682
; CURRENT FILING DATE: 2005-02-01
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 2403
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Plasmid pMHas5
US-10-523-682-1

Query Match 22.5%; Score 1166.2; DB 7; Length 2403;

Best Local Similarity 99.3%; Pred. No. 6.8e-267;

Matches 1171; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy	3370	CAAGCTTTCAGCTGCGCAAGCACTCAGGGCGCAAGGGCTGCTTAAGGAAGCGGAACACG	3429
Db	8	CTAGCTTTCAGCTGCGCAAGCACTCAGGGCGCAAGGGCTGCTTAAGGAAGCGGAACACG	67

Db 4160 CTGGGCAACACAGCAATCGGCTGCTCTGATGCGCGCGTGTTCGGCTGTTCAGCGCAGGG 4101
Qy 3855 GCGCCCGGTTCTTTTGTCAAGACCGACTGTTCGGTCCCTGATGAATGAATGCGAGGACGA 3914
Db 4100 GCGCCCGGTTCTTTTGTCAAGACCGACTGTTCGGTCCCTGATGAATGAATGCGAGGACGA 4041
Qy 3915 GCGACGCGGCTATCGTGGCTGGCCACACGCGGCGTTCCTTGGCAGCTGTGCTCGAGCT 3974
Db 4040 GCGACGCGGCTATCGTGGCTGGCCACACGCGGCGTTCCTTGGCAGCTGTGCTCGAGCT 3981
Qy 3975 TGTCACTGAAGCGGAGGAGTGGCTGCTATTTGGGCGAAGTGGCGGCGCAGGATCTCCT 4034
Db 3980 TGTCACTGAAGCGGAGGAGTGGCTGCTATTTGGGCGAAGTGGCGGCGCAGGATCTCCT 3921
Qy 4035 GTCACTCACTTGTCTCTGCGGAGAAAGTATCCATCATGCTGATGCAATGCGCGGCT 4094
Db 3920 GTCACTCACTTGTCTCTGCGGAGAAAGTATCCATCATGCTGATGCAATGCGCGGCT 3861
Qy 4095 GCATACGCTTGTATCGGCTGACCGACGCGGCGTTCCTTGGCAGCTGTGCTCGAGCT 4154
Db 3860 GCATACGCTTGTATCGGCTGACCGACGCGGCTTCCTTGGCAGCTGTGCTCGAGCT 3801
Qy 4155 AGCAGCTACTCGGATGGAAGCGGCTTGTGCTGATCAGGATGATCTGGACGAAAGCATCA 4214
Db 3800 AGCAGCTACTCGGATGGAAGCGGCTTGTGCTGATCAGGATGATCTGGACGAAAGCATCA 3741
Qy 4215 GGGGCTCGGCGCAGCGAACTGTTGCGCAGGCTCAAGGCGGCGAT - CCGGACGCGGAGGA 4274
Db 3740 GGGGCTCGGCGCAGCGAACTGTTGCGCAGGCTCAAGGCGGCGAT - CCGGACGCGGAGGA 3682
Qy 4275 TCTGCTGTGACCAATGCGGATGCTGCTGCGGCAATATCATGTTGGAATAATGCGCGCTT 4334
Db 3681 TCTGCTGTGACCAATGCGGATGCTGCTGCGGCAATATCATGTTGGAATAATGCGCGCTT 3622
Qy 4335 TTCTGGATTATCGACTGTGGCGGCTGGGTGTGCGGACCGCTATCAGGACATAGCGTT 4394
Db 3621 TTCTGGATTATCGACTGTGGCGGCTGGGTGTGCGGACCGCTATCAGGACATAGCGTT 3562
Qy 4395 GCGTACCGCTGATATGCTGGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCGTGTCT 4454
Db 3561 GCGTACCGCTGATATGCTGGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCGTGTCT 3502
Qy 4455 TTACGGTATCGCGCTCCCGATTCGACGCGATCGCCTTCTATCGCCTTCTTGACGAGTT 4514
Db 3501 TTACGGTATCGCGCTCCCGATTCGACGCGATCGCCTTCTATCGCCTTCTTGACGAGTT 3442
Qy 4515 CTCTGAGCGGACTCTCGGGTTTCGATA 4542
Db 3441 CTCTGAGCGGACTCTCGGGTTTCGAAA 3414

RESULT 7
US-11-082-154A-118/c
; Sequence 118, Application US/11082154A
; Publication No. US20060024820A1
; GENERAL INFORMATION:
; APPLICANT: Perkins, Edward
; APPLICANT: Perez, Carl
; APPLICANT: Lindenbaum, Michael
; APPLICANT: Greene, Amy
; APPLICANT: Leung, Josephine
; APPLICANT: Fleming, Elena
; APPLICANT: Stewart, Sandra
; APPLICANT: Shellard, Joan
; TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS
; FILE REFERENCE: 17084-022003 (420C)
; CURRENT APPLICATION NUMBER: US/11/082,154A
; CURRENT FILING DATE: 2005-03-15
; PRIOR APPLICATION NUMBER: 60/294,758
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: 60/366,891
; PRIOR FILING DATE: 2002-03-21

; PRIOR APPLICATION NUMBER: US 10/161,403
; PRIOR FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 118
; LENGTH: 17384
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pFK161 Plasmid
US-11-082-154A-118

Query Match 22.3%; Score 1154.4; DB 12; Length 17384;
Best Local Similarity 99.8%; Pred. No. 1.2e-263;
Matches 1166; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 3375 TTCAACGCTGCGCAGCACTCAGGGCGCAAGGGCTGCTAAAGGAAGCGGAACACGTAGAA 3434
Db 2845 TTCAACGCTGCGCAGCACTCAGGGCGCAAGGGCTGCTAAAGGAAGCGGAACACGTAGAA 2786
Qy 3435 AGCCAGTCCGCGAAGAAACGGTGTGACCCCGATGAATGTCTAGCTACTGGGCTATCTGGAC 3494
Db 2785 AGCCAGTCCGCGAAGAAACGGTGTGACCCCGATGAATGTCTAGCTACTGGGCTATCTGGAC 2726
Qy 3495 AAGGGAACCGCAAGCCCAAGAGAAAGCAGTAGCTTGCAGTGGGCTTACATGGCGATA 3554
Db 2725 AAGGGAACCGCAAGCCCAAGAGAAAGCAGTAGCTTGCAGTGGGCTTACATGGCGATA 2666
Qy 3555 GCTAGACTGGGCGGTTTTATGGACAGCAAGCGAAACCGGAATTGCCAGTGGGCGCCCTC 3614
Db 2665 GCTAGACTGGGCGGTTTTATGGACAGCAAGCGAAACCGGAATTGCCAGTGGGCGCCCTC 2606
Qy 3615 TGGTAAGGTTGGGAAGCCCTGCAAGATGAATCTGATGCTCTTCTTGGCCGCAAGGATCTG 3674
Db 2605 TGGTAAGGTTGGGAAGCCCTGCAAGATGAATCTGATGCTCTTCTTGGCCGCAAGGATCTG 2546
Qy 3675 ATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCGTTTTCGCATGATGA 3734
Db 2545 ATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCGTTTTCGCATGATGA 2486
Qy 3735 ACAAGATGGAATTCACACAGGTTCTTCGGCGCTTGGGTGGAGAGGCTATTCGGCTATGA 3794
Db 2485 ACAAGATGGAATTCACACAGGTTCTTCGGCGCTTGGGTGGAGAGGCTATTCGGCTATGA 2426
Qy 3795 CTGGGCAACACAGCAATCGGCTGCTGATGCGCGCTGTTCCGGCTGTTCAGCGCAGGG 3854
Db 2425 CTGGGCAACACAGCAATCGGCTGCTGATGCGCGCTGTTCCGGCTGTTCAGCGCAGGG 2366
Qy 3855 GCGCCCGGTTCTTTTGTCAAGACCGGCTGTCCGGTGCCTGATGAATGAATGCACTGCAAGGACGA 3914
Db 2365 GCGCCCGGTTCTTTTGTCAAGACCGGCTGTCCGGTGCCTGATGAATGCACTGCAAGGACGA 2306
Qy 3915 GGCAGCGCGCTATCGTGGCTGGCCACGACGGGCGTTCTTTCGCGAGCTGTGCTCGACGT 3974
Db 2305 GGCAGCGCGCTATCGTGGCTGGCCACGACGGGCGTTCTTTCGCGAGCTGTGCTCGACGT 2246
Qy 3975 TGTCACTGAAGCGGAGGAGCTGGCTGCTATTTGGGCGAAGTGGCGGCGAGGATCTCCT 4034
Db 2245 TGTCACTGAAGCGGAGGAGCTGGCTGCTATTTGGGCGAAGTGGCGGCGAGGATCTCCT 2186
Qy 4035 GTCACTCACCTTCTCTGCGGAGAAAGTATCCATCATGCTGATGCAATGCGCGGCT 4094
Db 2185 GTCACTCACCTTCTCTGCGGAGAAAGTATCCATCATGCTGATGCAATGCGCGGCT 2126
Qy 4095 GCATACGCTTGTATCGGCTACCTTGGCCATTCGACCAACAGCGGAACATCGCATCGAGCG 4154
Db 2125 GCATACGCTTGTATCGGCTACCTTGGCCATTCGACCAACAGCGGAACATCGCATCGAGCG 2066
Qy 4155 AGCAGCTACTCGGATGGAAGCGGCTTGTGCTGATCAGGATGATCTGGACGAAAGCATCA 4214
Db 2065 AGCAGCTACTCGGATGGAAGCGGCTTGTGCTGATCAGGATGATCTGGACGAAAGCATCA 2006
Qy 4215 GGGGCTCGCGCAGCGAACTGTTTCGCGCAGGCTCAAGGCGCGCATGCTGCCGCGCAGGA 4274

2005	Db		GGGGCTCGCGCCAGCCGAACTGTTTCGCAGGCTCAAGGCGCGCAT-CCCGACGGCGAGGA	1947
4275	Qy		TCTCGTGTGACCCATGGCGCATGCTCTTTCGCCGAATATCATGGTGGAAAAATGGCCGCCTT	4334
1946	Db		TCTCGTGTGACCCATGGCGCATGCTCTTTCGCCGAATATCATGGTGGAAAAATGGCCGCCTT	1887
4335	Qy		TTCTGGAATTCATCGACTGTGGCGCGCTGGGTGTGGCGGACCGCTATCAGAGCATAGCGTT	4394
1886	Db		TTCTGGAATTCATCGACTGTGGCGCGCTGGGTGTGGCGGACCGCTATCAGAGCATAGCGTT	1827
4395	Qy		GGCTACCCGTGATATTGCTGAAGAGCTTGGCGCGGAATGGGCTGACCGCTTCTCTGCTGCT	4454
1826	Db		GGCTACCCGTGATATTGCTGAAGAGCTTGGCGCGGAATGGGCTGACCGCTTCTCTGCTGCT	1767
4455	Qy		TTACGGTATCGCGCTCCCGAATTTCGACAGCGCATCGGCTTCTATCGGCTTCTTCGACAGATT	4514
1766	Db		TTACGGTATCGCGCTCCCGAATTTCGACAGCGCATCGGCTTCTATCGGCTTCTTCGACAGATT	1707
4515	Qy		CTTCTGAGCGGAGACTCTGGGGTTCGATA	4542
1706	Db		CTTCTGAGCGGAGACTCTGGGGTTCGAAA	1679

RESULT 8

```

US-10-525-710-66
; Sequence 66, Application US/10525710
; Publication No. US20050260721A1
; GENERAL INFORMATION:
; APPLICANT: Kroger, Burkhard
; APPLICANT: Zelder, Oskar
; APPLICANT: Kolpprogge, Corinna
; APPLICANT: Schroder, Hartwig
; APPLICANT: Hafner, Stefan
; TITLE OF INVENTION: Method for Zymotic Production of Fine Chemicals Containing
; TITLE OF INVENTION: Sulphur (mety)
; FILE REFERENCE: 13111-00006-US
; CURRENT APPLICATION NUMBER: US/10/525,710
; CURRENT FILING DATE: 2005-02-24
; PRIOR APPLICATION NUMBER: PCT/EP 2003/009453
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: DE 102 39 082.7
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 66
; LENGTH: 4323
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence:plasmid
US-10-525-710-66

```

RESULT 9

```

US-10-525-674-58
; Sequence 58, Application US/10525674
; Publication No. US20060003425A1
; GENERAL INFORMATION:
; APPLICANT: Kroger, Burkhard
; APPLICANT: Zelder, Oskar
; APPLICANT: Kolpprogge, Corinna
; APPLICANT: Schroder, Hartwig
; APPLICANT: Hafner, Stefan
; TITLE OF INVENTION: Method for Zymotic Production of Fine Chemicals Containing
; TITLE OF INVENTION: Sulphur (Meta)

```

APPLICANT:	hainer, steran
TITLE OF INVENTION:	Method for Zymotic Production of Fine Chemicals Containing Sulphur (Meta)

```
; FILE REFERENCE: 13111-00002-US
; CURRENT APPLICATION NUMBER: US/10/525,674
; CURRENT FILING DATE: 2005-02-24
; PRIOR APPLICATION NUMBER: PCT/EP 2003/009452
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: DE 102 39 073.8
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 58
; LENGTH: 4323
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: plasmid
US-10-525-674-58

Query Match      22.0%; Score 1136.4; DB 8; Length 4323;
Best Local Similarity 99.9%; Pred. No. 1.1e-259;
Matches 1137; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3405 GGGCTGCTAAAGGAAGCGGAACACGTTAGAAAGCCAGTCCGCAGAAACCGGTGCTGACCCCG 3464
Db 1359 GGGCTGCTAAAGGAAGCGGAACACGTTAGAAAGCCAGTCCGCAGAAACCGGTGCTGACCCCG 1418

Qy 3465 GATGAATGTCAGTACTGGGCTATCTGCAGCAAGGGAACCGCAAGCGCAAGAGAAAGCA 3524
Db 1419 GATGAATGTCAGTACTGGGCTATCTGCAGCAAGGGAACCGCAAGCGCAAGAGAAAGCA 1478

Qy 3525 GGTAGCTTGCACTGGGCTTACATGCGCATAGCTAGACTGGGCGGTTTTATGACAGCAAG 3584
Db 1479 GGTAGCTTGCACTGGGCTTACATGCGCATAGCTAGACTGGGCGGTTTTATGACAGCAAG 1538

Qy 3585 CGAACCGGAATGCGCAGTGGGGGCCCTCTGGTAAAGTTGGGAAGCCCTCAAGTAA 3644
Db 1539 CGAACCGGAATGCGCAGTGGGGGCCCTCTGGTAAAGTTGGGAAGCCCTCAAGTAA 1598

Qy 3645 CTGGATGCTTTCTTGCGCCCAAGATCTGATGGCGAGGGGATCAAGATCTGATCAAGA 3704
Db 1599 CTGGATGCTTTCTTGCGCCCAAGATCTGATGGCGAGGGGATCAAGATCTGATCAAGA 1658

Qy 3705 GACAGGATGAGGATCGTTTCCGATGATTAACAAAGATGATGACGAGGTTCTCCGGC 3764
Db 1659 GACAGGATGAGGATCGTTTCCGATGATTAACAAAGATGATGACGAGGTTCTCCGGC 1718

Qy 3765 CGCTTGGTGGAGAGGCTATTGCGCTATGATCTGGGCAACAGCAATCGGCTGCTTGA 3824
Db 1719 CGCTTGGTGGAGAGGCTATTGCGCTATGATCTGGGCAACAGCAATCGGCTGCTTGA 1778

Qy 3825 TGCCGCGCTGTTCCGGCTGTCAGCGCAGGGCGCCGGTCTTTTGTGTCAGAGCGACCT 3884
Db 1779 TGCCGCGCTGTTCCGGCTGTCAGCGCAGGGCGCCGGTCTTTTGTGTCAGAGCGACCT 1838

Qy 3885 GTCCGGTGCCTGAATGAATGACGAGCAGCGCGGCTATCGTGGCTGGCCACGAC 3944
Db 1839 GTCCGGTGCCTGAATGAATGACGAGCAGCGCGGCTATCGTGGCTGGCCACGAC 1898

Qy 3945 GGGCGCTTCTTGGCGAGCTGCTCGACGTTGTCTACTGAAGCGGGAAGGACTGCTGCT 4004
Db 1899 GGGCGCTTCTTGGCGAGCTGCTCGACGTTGTCTACTGAAGCGGGAAGGACTGCTGCT 1958

Qy 4005 ATTGGGCGAAGTGGCGGGCAGGATCTCTGTGTCATCTCACTTGTCTCTCGGAGAAAT 4064
Db 1959 ATTGGGCGAAGTGGCGGGCAGGATCTCTGTGTCATCTCACTTGTCTCTCGGAGAAAT 2018

Qy 4065 ATCCATCATGGCTGATGAATGCGCGGCTGATACGCTTGATCCGGCTACTGCGCCATT 4124
Db 2019 ATCCATCATGGCTGATGAATGCGCGGCTGATACGCTTGATCCGGCTACTGCGCCATT 2078

Qy 4125 CGACACCAAGCGAAACATCGCATCGAGCGAGCAGTACTCGGATGGAAGCGGCTTTGT 4184
Db 2079 CGACACCAAGCGAAACATCGCATCGAGCGAGCAGTACTCGGATGGAAGCGGCTTTGT 2138

; FILE REFERENCE: 13111-00006-US
; CURRENT APPLICATION NUMBER: US/10/525,710
; PUBLICATON NO. US20050260721A1
; GENERAL INFORMATION:
; APPLICANT: Kroger, Burkhard
; APPLICANT: Zelder, Oskar
; APPLICANT: Kolpprogge, Corinna
; APPLICANT: Schroder, Hartwig
; APPLICANT: Hafner, Stefan
; TITLE OF INVENTION: Method for Zymotic Production of Fine Chemicals Containing
; TITLE OF INVENTION: Sulphur (mety)
; FILE REFERENCE: 13111-00006-US
; CURRENT APPLICATION NUMBER: US/10/525,710
; CURRENT FILING DATE: 2005-02-24
; PRIOR APPLICATION NUMBER: PCT/EP 2003/009453
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: DE 102 39 082.7
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 65
; LENGTH: 5091
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: plasmid
US-10-525-710-65

Query Match      22.0%; Score 1136.4; DB 8; Length 5091;
Best Local Similarity 99.9%; Pred. No. 1.2e-259;
Matches 1137; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3405 GGGCTGCTAAAGGAAGCGGAACACGTTAGAAAGCCAGTCCGCAGAAACCGGTGCTGACCCCG 3464
Db 358 GGGCTGCTAAAGGAAGCGGAACACGTTAGAAAGCCAGTCCGCAGAAACCGGTGCTGACCCCG 417

Qy 3465 GATGAATGTCAGTACTGGGCTATCTGCAGCAAGGGAACCGCAAGCGCAAGAGAAAGCA 3524
Db 418 GATGAATGTCAGTACTGGGCTATCTGCAGCAAGGGAACCGCAAGCGCAAGAGAAAGCA 477

Qy 3525 GGTAGCTTGCACTGGGCTTACATGCGCATAGCTAGACTGGGCGGTTTTATGACAGCAAG 3584
Db 478 GGTAGCTTGCACTGGGCTTACATGCGCATAGCTAGACTGGGCGGTTTTATGACAGCAAG 537

Qy 3585 CGAACCGGAATGCGCAGTGGGGGCCCTCTGGTAAAGTTGGGAAGCCCTCGAAAGTAA 3644
Db 538 CGAACCGGAATGCGCAGTGGGGGCCCTCTGGTAAAGTTGGGAAGCCCTCGAAAGTAA 597
```

3645 CTGGATGCGTTCTTTCGCCCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGA 3704
598 CTGGATGCGTTCTTTCGCCCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGA 657
3705 GACAGGATGAGGATCGTTTTCGCATGATTGAACAGATGATTGACGCGAGGTTCTCCGGC 3764
658 GACAGGATGAGGATCGTTTTCGCATGATTGAACAGATGATTGACGCGAGGTTCTCCGGC 717
3765 CGCTTGGGTGAGAGGCTATTGCGCTATGATCTGGGCAACAAGCAATCGGCTGCTCTGA 3824
718 CGCTTGGGTGAGAGGCTATTGCGCTATGATCTGGGCAACAAGCAATCGGCTGCTCTGA 777
3825 TGCCGCGCTGTTCCGGCTGTACGCCAGGGGGGCCCGGTTCTTTTGTCAAGACCGACCT 3884
778 TGCCGCGCTGTTCCGGCTGTACGCCAGGGGGGCCCGGTTCTTTTGTCAAGACCGACCT 837
3885 GTCCGCGTCCCTGATGATGAATGACGACGACGCGGCTATCGTGGCTGGCCACGAC 3944
838 GTCCGCGTCCCTGATGATGAATGACGACGACGCGGCTATCGTGGCTGGCCACGAC 897
3945 GGGCGTTCTTTCGCCGAGCTGTCTCGAGCTGTGTCTCACTGAAGCGGGAAGGACTGTGCTGCT 4004
898 GGGCGTTCTTTCGCCGAGCTGTCTCGAGCTGTGTCTCACTGAAGCGGGAAGGACTGTGCTGCT 957
4005 ATTGGGCGAAGTGCCGGGCGAGGATCTCTGTCTATCTCACTTGTCTCTGCGGAGAAAGT 4064
958 ATTGGGCGAAGTGCCGGGCGAGGATCTCTGTCTATCTCACTTGTCTCTGCGGAGAAAGT 1017
4065 ATCCATCATGCTGATGCAATGCGGCGCTGCATACGCTTGTATCGGCTACCTGCCCAAT 4124
1018 ATCCATCATGCTGATGCAATGCGGCGCTGCATACGCTTGTATCGGCTACCTGCCCAAT 1077
4125 CGACCAACGAGCAAAATCGCATCGAGGACGACGTAATCGGATGGAAGCGGCTTGT 4184
1078 CGACCAACGAGCAAAATCGCATCGAGGACGACGTAATCGGATGGAAGCGGCTTGT 1137
4185 CGATCAGGATGATCTGGACGAGAGCATCAGGGGCTCGCGCAGCGCAACTGTTTGGCCAG 4244
1138 CGATCAGGATGATCTGGACGAGAGCATCAGGGGCTCGCGCAGCGCAACTGTTTGGCCAG 1197
4245 GCTCAAGCGGCGCATGCCGAGCGGAGGATCTGCTGTGACCCATGCGGATGCTTGT 4304
1198 GCTCAAGCGGCGCATGCCGAGCGGAGGATCTGCTGTGACCCATGCGGATGCTTGT 1257
4305 GCCGAATATCATGTTGGAAATGCGGCTTCTTCTGATTCATCGACTGTGCGCGCTGGG 4364
1258 GCCGAATATCATGTTGGAAATGCGGCTTCTTCTGATTCATCGACTGTGCGCGCTGGG 1317
4365 TGTGCGGACCGCTATCAGGACATAGCGTTGGCTACCGCTGATATTGCTGAAGAGCTTGG 4424
1318 TGTGCGGACCGCTATCAGGACATAGCGTTGGCTACCGCTGATATTGCTGAAGAGCTTGG 1377
4425 CGCGAATGGGCTGACCGCTTCTCGTCTTTACGGATTCGCGCTCCCGATTCGCGAGCG 4484
1378 CGCGAATGGGCTGACCGCTTCTCGTCTTTACGGATTCGCGCTCCCGATTCGCGAGCG 1437
4485 CATGCGCTTCTATCGCTTCTTACGAGTCTTCTGAGCGGAGCTCTGCGGTTCTGATA 4542
1438 CATGCGCTTCTATCGCTTCTTACGAGTCTTCTGAGCGGAGCTCTGCGGTTCTGATA 1495

RESULT 11

US-10-525-674-57
; Sequence 57, Application US/10525674
; Publication No. US20060003425A1
; GENERAL INFORMATION:
; APPLICANT: Kroger, Burkhard
; APPLICANT: Zelder, Oskar
; APPLICANT: Kolprogge, Corinna
; APPLICANT: Schroder, Hartwig
; APPLICANT: Hafner, Stefan
; TITLE OF INVENTION: Method for Zymotic Production of Fine Chemicals Containing

; TITLE OF INVENTION: Sulphur (Meta)
; FILE REFERENCE: 13111-00002-US
; CURRENT APPLICATION NUMBER: US/10/525,674
; CURRENT FILING DATE: 2005-02-24
; PRIOR APPLICATION NUMBER: PCT/EP 2003/009452
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: DE 102 39 073.8
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 57
; LENGTH: 5091
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: plasmid
US-10-525-674-57

Query Match 22.0%; Score 1136.4; DB 8; Length 5091;
Best Local Similarity .99.9%; Pred. No. 1.2e-259;
Matches 1137; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3405 GGGCTGCTAAAGGAAGCGGAACACGTAGAAAGCCAGTCCGCGAAGAACGGTCTGCACCCCG 3464
DB 358 GGGCTGCTAAAGGAAGCGGAACACGTAGAAAGCCAGTCCGCGAAGAACGGTCTGCACCCCG 417
QY 3465 GATGAATGTCTCAGCTACTGCGCTATCTGGACAAAGGAAACCGCAAGCGCAAGAGCA 3524
DB 418 GATGAATGTCTCAGCTACTGCGCTATCTGGACAAAGGAAACCGCAAGCGCAAGAGCA 477
QY 3525 GGTAGCTTGCAGTGGGCTTACATGGCGATAGCTAGACTGGGCGGTTTTATGGACAGCAAG 3584
DB 478 GGTAGCTTGCAGTGGGCTTACATGGCGATAGCTAGACTGGGCGGTTTTATGGACAGCAAG 537
QY 3585 CGAACCCGGAATGCGCAGTGGGCGGCTCTGTTAAGGTTGGGAAGCCCTCCAAAGTAAA 3644
DB 538 CGAACCCGGAATGCGCAGTGGGCGGCTCTGTTAAGGTTGGGAAGCCCTCCAAAGTAAA 597
QY 3645 CTGGATGCGTTCTTTCGCCCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGA 3704
DB 598 CTGGATGCGTTCTTTCGCCCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGA 657
QY 3705 GACAGGATGAGGATCGTTTTCGCATGATTGAACAGATGATTGACGCGAGGTTCTCCGGC 3764
DB 658 GACAGGATGAGGATCGTTTTCGCATGATTGAACAGATGATTGACGCGAGGTTCTCCGGC 717
QY 3765 CGCTTGGGTGAGAGGCTATTGCGCTATGATCTGGGCAACAAGCAATCGGCTGCTCTGA 3824
DB 718 CGCTTGGGTGAGAGGCTATTGCGCTATGATCTGGGCAACAAGCAATCGGCTGCTCTGA 777
QY 3825 TGCCGCGCTGTTCCGGCTGTACGCCAGGGGGGCCCGGTTCTTTTGTCAAGACCGACCT 3884
DB 778 TGCCGCGCTGTTCCGGCTGTACGCCAGGGGGGCCCGGTTCTTTTGTCAAGACCGACCT 837
QY 3885 GTCCGCGCTGTTCCGGCTGTACGCCAGGGGGGCCCGGTTCTTTTGTCAAGACCGACCT 3944
DB 838 GTCCGCGCTGTTCCGGCTGTACGCCAGGGGGGCCCGGTTCTTTTGTCAAGACCGACCT 897
QY 3945 GGGGCTTCTTTCGCCGAGCTGTCTCGAGCTGTGTCTCACTGAAGCGGGAAGGACTGTGCTGCT 4004
DB 898 GGGGCTTCTTTCGCCGAGCTGTCTCGAGCTGTGTCTCACTGAAGCGGGAAGGACTGTGCTGCT 957
QY 4005 ATTGGGCGAAGTGCCGGGCGAGGATCTCTGTCTATCTCACTTGTATCGGCTACCTGCCCAAT 4064
DB 958 ATTGGGCGAAGTGCCGGGCGAGGATCTCTGTCTATCTCACTTGTATCGGCTACCTGCCCAAT 1017
QY 4065 ATCCATCATGCTGATGCAATGCGGCGCTGCATACGCTTGTATCGGCTACCTGCCCAAT 4124
DB 1018 ATCCATCATGCTGATGCAATGCGGCGCTGCATACGCTTGTATCGGCTACCTGCCCAAT 1077
QY 4125 CGACCAACGAGCAAAATCGCATCGAGGACGACGTAATCGGATGGAAGCGGCTTGT 4184
DB 1078 CGACCAACGAGCAAAATCGCATCGAGGACGACGTAATCGGATGGAAGCGGCTTGT 1137

QY 4185 CGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCAGCCGAACTGTTTCGCCAG 4244
Db 1138 CGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCAGCCGAACTGTTTCGCCAG 1197
QY 4245 GCTCAAGCGCGCATGCCCGACCGGAGGATCTCGTCTGTGACCCCATGSCGATGCTGCTT 4304
Db 1198 GCTCAAGCGCGCATGCCCGACCGGAGGATCTCGTCTGTGACCCCATGSCGATGCTGCTT 1257
QY 4305 GCCGAATATCATGGTGGAAATGGCCGCTTTCTGGATTCATCGACTGTGCGCCGGCTGGG 4364
Db 1258 GCCGAATATCATGGTGGAAATGGCCGCTTTCTGGATTCATCGACTGTGCGCCGGCTGGG 1317
QY 4365 TGTGCGGACCGCTATCAGGACATAGCTTGGCTACCCGCTGATATTGCTGAAGAGCTTGG 4424
Db 1318 TGTGCGGACCGCTATCAGGACATAGCTTGGCTACCCGCTGATATTGCTGAAGAGCTTGG 1377
QY 4425 CGCGGAATGGGCTGACCGCTTCTCGTGTCTTTACGGTATCGCCGCTCCCGATTCGCGAGCG 4484
Db 1378 CGCGGAATGGGCTGACCGCTTCTCGTGTCTTTACGGTATCGCCGCTCCCGATTCGCGAGCG 1437
QY 4485 CATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGAGCTCTGGGTTTCGATA 4542
Db 1438 CATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGAGCTCTGGGTTTCGATA 1495

RESULT 12
US-10-525-710-69
; Sequence 69, Application US/10525710
; Publication No. US20050260721A1
; GENERAL INFORMATION:
; APPLICANT: Kroger, Burkhard
; APPLICANT: Zelder, Oskar
; APPLICANT: Kolpprogge, Corinna
; APPLICANT: Schroder, Hartwig
; TITLE OF INVENTION: Method for Zymotic Production of Fine Chemicals Containing
; TITLE OF INVENTION: Sulphur (mety)
; FILE REFERENCE: 13111-00006-US
; CURRENT APPLICATION NUMBER: US/10/525,710
; CURRENT FILING DATE: 2005-02-24
; PRIOR APPLICATION NUMBER: PCT/EP 2003/009453
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: DE 102 39 082.7
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 69
; LENGTH: 5860
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURES:
; OTHER INFORMATION: Description of the artificial sequence: plasmid
US-10-525-710-69

Query Match 22.0%; Score 1136.4; DB 8; Length 5860;
Best Local Similarity 99.9%; Pred. No. 1.3e-259;
Matches 1137; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3405 GGGCTGCTAAAGGAAGCGGAACACCGTAGAAGCCAGTCCGAGAAACCGGTGCTGACCCCG 3464
Db 1652 GGGCTGCTAAAGGAAGCGGAACACCGTAGAAGCCAGTCCGAGAAACCGGTGCTGACCCCG 1711
QY 3465 GATGAATGTACGTTACTGGGCTATCTGGACAAAGGAAACCGCAAGCAAGCAAGCA 3524
Db 1712 GATGAATGTACGTTACTGGGCTATCTGGACAAAGGAAACCGCAAGCAAGCAAGCA 1771
QY 3525 GGTAGCTTGCAGTGGGCTTACATGCGGATAGCTAGCTGGGCTTTTATGACAGCAAG 3584
Db 1772 GGTAGCTTGCAGTGGGCTTACATGCGGATAGCTAGCTGGGCTTTTATGACAGCAAG 1831
QY 3585 CGAACCGGAATGCCAGCTGGGGGCCCTCTGTTAAGGTTGGGAAGCCCTCCAAAGTAAA 3644

Db 1832 CGAACCGGAATGCCAGCTGGGGCGCCCTCTGTGTAAAGTTGGGAAGCCCTCCAAAGTAAA 1891
QY 3645 CTGGATGGCTTTCTTGGCCGCAAGGATCTGATGGCGAGGGGATCAAGATCTGATCAAGA 3704
Db 1892 CTGGATGGCTTTCTTGGCCGCAAGGATCTGATGGCGAGGGGATCAAGATCTGATCAAGA 1951
QY 3705 GACAGGATGAGGATCGTTTTCGCATGATTGAACAAGATGGATTGCACGAGGTTCTCGGGC 3764
Db 1952 GACAGGATGAGGATCGTTTTCGCATGATTGAACAAGATGGATTGCACGAGGTTCTCGGGC 2011
QY 3765 CGCTTGGGTGGAAGGCTATTCGGCTATGACTGGGCAACAAGATTCGGCTCTGTA 3824
Db 2012 CGCTTGGGTGGAAGGCTATTCGGCTATGACTGGGCAACAAGATTCGGCTCTGTA 2071
QY 3825 TGGCGCCGCTTTCGGGCTGTTCAGGCGAGGGGCGCCCGGTTCTTTTCTCAAGACCGACCT 3884
Db 2072 TGGCGCCGCTTTCGGGCTGTTCAGGCGAGGGGCGCCCGGTTCTTTTCTCAAGACCGACCT 2131
QY 3885 GTCCGGTGCCTCTGAATGAACCTGCAGGACGAGGCGAGCGGCTATCGTGGCTGGCCACGAC 3944
Db 2132 GTCCGGTGCCTCTGAATGAACCTGCAGGACGAGGCGAGCGGCTATCGTGGCTGGCCACGAC 2191
QY 3945 GGGGCTTCTTGGCGAGCTGTGCTGACGTTGTGTCATGAAAGCGGGAAGGAGCTGGCTGCT 4004
Db 2192 GGGGCTTCTTGGCGAGCTGTGCTGACGTTGTGTCATGAAAGCGGGAAGGAGCTGGCTGCT 2251
QY 4005 ATTGGGCGAAGTCCCGGGCAGGATCTCCTGTGTCATCTCAGCTTGTCTCTCCGAGAAAGT 4064
Db 2252 ATTGGGCGAAGTCCCGGGCAGGATCTCCTGTGTCATCTCAGCTTGTCTCTCCGAGAAAGT 2311
QY 4065 ATCCATCATGGCTGATGCAATGGCGGCTGCATACGCTTGTGATCCGGCTACCTGCCCAT 4124
Db 2312 ATCCATCATGGCTGATGCAATGGCGGCTGCATACGCTTGTGATCCGGCTACCTGCCCAT 2371
QY 4125 CGACCAACGAAGCAACATCGCATCGAGCGAGCAGTACTCGGATGGAAGCGGCTCTGT 4184
Db 2372 CGACCAACGAAGCAACATCGCATCGAGCGAGCAGTACTCGGATGGAAGCGGCTCTGT 2431
QY 4185 CGATCAGGATGATCTGGACGAAGCATCAGGGGCTCGCGCCAGCCGAACTGTTTCGCCAG 4244
Db 2432 CGATCAGGATGATCTGGACGAAGCATCAGGGGCTCGCGCCAGCCGAACTGTTTCGCCAG 2491
QY 4245 GCTCAAGCGCGCATGCCCGACCGAGGATCTGCTGTGACCCCATGGCGATGCTGCTT 4304
Db 2492 GCTCAAGCGCGCATGCCCGACCGAGGATCTGCTGTGACCCCATGGCGATGCTGCTT 2551
QY 4305 GCCGAATATCATGGTGGAAATGGCCGCTTTCTGGATTCATCGACTGTGCGCCGGCTGGG 4364
Db 2552 GCCGAATATCATGGTGGAAATGGCCGCTTTCTGGATTCATCGACTGTGCGCCGGCTGGG 2611
QY 4365 TGTGGCGGACCGCTATCAGGACATAGGTTGGCTTACCCGCTGATATTGCTGAAGAGCTGG 4424
Db 2612 TGTGGCGGACCGCTATCAGGACATAGGTTGGCTTACCCGCTGATATTGCTGAAGAGCTGG 2671
QY 4425 CGCGGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGCCGCTCCCGATTCGACGCG 4484
Db 2672 CGCGGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGCCGCTCCCGATTCGACGCG 2731
QY 4485 CATCGCCTTCTATCGCCTTCTTGAAGGTTCTTCTGAGCGGAGCTCTGGGTTTCGATA 4542
Db 2732 CATCGCCTTCTATCGCCTTCTTGAAGGTTCTTCTGAGCGGAGCTCTGGGTTTCGATA 2789

RESULT 13
US-10-525-710-74
; Sequence 74, Application US/10525710
; Publication No. US20050260721A1
; GENERAL INFORMATION:
; APPLICANT: Kroger, Burkhard
; APPLICANT: Zelder, Oskar
; APPLICANT: Kolpprogge, Corinna
; APPLICANT: Schroder, Hartwig
; APPLICANT: Hafner, Stefan

; TITLE OF INVENTION: Method for Zymotic Production of Fine Chemicals Containing
; FILE REFERENCE: Sulphur (mef)
; CURRENT APPLICATION NUMBER: US/10/525,710
; PRIOR FILING DATE: 2005-02-24
; PRIOR APPLICATION NUMBER: PCT/EP 2003/009453
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: DE 102 39 082.7
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 74
; LENGTH: 5860
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: plasmid
US-10-525-710-74

Query Match 22.0%; Score 1136.4; DB 8; Length 5860;
Best Local Similarity 99.9%; Pred. No. 1.3e-259;
Matches 1137; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3405 GGGCTGCTAAAGGAGCGGACACCTAGAAAGCCAGTCCGAGAAACGGTGTGACCCCG 3464
DB 1652 GGGCTGCTAAAGGAGCGGACACCTAGAAAGCCAGTCCGAGAAACGGTGTGACCCCG 1711

QY 3465 GATGAATGTACGTACTGGGCTATCTGGACAAAGGAGGAAACGCAAGCGCAAGAGAAAGCA 3524
DB 1712 GATGAATGTACGTACTGGGCTATCTGGACAAAGGAGGAAACGCAAGCGCAAGAGAAAGCA 1771

QY 3525 GGTAGCTTCGATGGGCTTACATGCGGATAGTACGCTGGGCGGTTTATGACAGCAAG 3584
DB 1772 GGTAGCTTCGATGGGCTTACATGCGGATAGTACGCTGGGCGGTTTATGACAGCAAG 1831

QY 3585 CGAACCGGAATGCCAGCTGGGCGCCCTCTGTAAGTTGGGAAGCCCTGCAAGATGAAA 3644
DB 1832 CGAACCGGAATGCCAGCTGGGCGCCCTCTGTAAGTTGGGAAGCCCTGCAAGATGAAA 1891

QY 3645 CTGGATGCTTCTTGGCCCAAGATCTGATGGCGAGGGGATCAAGATCTGATCAAGA 3704
DB 1892 CTGGATGCTTCTTGGCCCAAGATCTGATGGCGAGGGGATCAAGATCTGATCAAGA 1951

QY 3705 GACAGGATGAGGATCGTTTCGATGATGAACAGATGATGACGAGGTTCTCGGC 3764
DB 1952 GACAGGATGAGGATCGTTTCGATGATGAACAGATGATGACGAGGTTCTCGGC 2011

QY 3765 CGCTTGGGTGAGAGGCTATTGCGCTATGATCGGCAACAGCAATCGGCTGCTGTA 3824
DB 2012 CGCTTGGGTGAGAGGCTATTGCGCTATGATCGGCAACAGCAATCGGCTGCTGTA 2071

QY 3825 TGCGCGCTGTTCCGCTGTGAGCGAGGGGCGCGGTTCTTTTGTCAAGACCGACCT 3884
DB 2072 TGCGCGCTGTTCCGCTGTGAGCGAGGGGCGCGGTTCTTTTGTCAAGACCGACCT 2131

QY 3885 GTCCGGTCCCTGATGAATGAACTGACGAGCGGCGGCTATCGTGGCTGGCCACGAC 3944
DB 2132 GTCCGGTCCCTGATGAATGAACTGACGAGCGGCGGCTATCGTGGCTGGCCACGAC 2191

QY 3945 GGGCGTTCTTGGCGAGCTGTCTGAGTTGTCTACTGAAGCGGGAAGGACTGGCTGCT 4004
DB 2192 GGGCGTTCTTGGCGAGCTGTCTGAGTTGTCTACTGAAGCGGGAAGGACTGGCTGCT 2251

QY 4005 ATTGGGCGAAGTCCGGGCGAGGATCTCTGTCATCTCACTTGTCTGCTGCGGAGAGT 4064
DB 2252 ATTGGGCGAAGTCCGGGCGAGGATCTCTGTCATCTCACTTGTCTGCTGCGGAGAGT 2311

QY 4065 ATCCATCATGGCTGATGCAATGCGGCGGCTGATACGCTTGTATCGGCTACTGCCCCATT 4124
DB 2312 ATCCATCATGGCTGATGCAATGCGGCGGCTGATACGCTTGTATCGGCTACTGCCCCATT 2371

QY 4125 CGACCAACGCAAGCAATCGCATCGAGCGAGCACTACTCGGATGGAAGCCGCTTTGT 4184

DB 2372 CGACCAACGCAAGCAATCGCATCGAGCGAGCACTACTCGGATGGAGCCGCTTTGT 2431
QY 4185 CGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCAGCCGAACTGTTGCGCAG 4244
DB 2432 CGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCAGCCGAACTGTTGCGCAG 2491

QY 4245 GCTCAAGCGCGCATGCCGACCGGAGGATCTCTGTGTGATCCCATGCGGATGCTGCTT 4304
DB 2492 GCTCAAGCGCGCATGCCGACCGGAGGATCTCTGTGTGATCCCATGCGGATGCTGCTT 2551

QY 4305 GCCGAATATCATGTGGAAATGCGCTTTCTTGGATTCTGACTGTGTGCGGCTGG 4364
DB 2552 GCCGAATATCATGTGGAAATGCGCTTTCTTGGATTCTGACTGTGTGCGGCTGG 2611

QY 4365 TGTGCGGACCGCTATCAGGACATAGCGTTGGCTACCGGTGATATGCTGAAGAGCTTGG 4424
DB 2612 TGTGCGGACCGCTATCAGGACATAGCGTTGGCTACCGGTGATATGCTGAAGAGCTTGG 2671

QY 4425 CGGCAATGGGCTGACCGCTTCTCTGTCTTACGGTATCGCGCTCCCGATTGCGCAGG 4484
DB 2672 CGGCAATGGGCTGACCGCTTCTCTGTCTTACGGTATCGCGCTCCCGATTGCGCAGG 2731

QY 4485 CATCGCTTCTATCGCTTCTTGACGAGTTCTTCTGAGCGGAGCTCTGGGCTTCGATA 4542
DB 2732 CATCGCTTCTATCGCTTCTTGACGAGTTCTTCTGAGCGGAGCTCTGGGCTTCGATA 2789

RESULT 14
US-10-525-674-61
; Sequence 61, Application US/10525674
; Publication No. US20060003425A1
; GENERAL INFORMATION:
; APPLICANT: Kroger, Burkhard
; APPLICANT: Zelder, Oskar
; APPLICANT: Kolpprogge, Corinna
; APPLICANT: Schroder, Hartwig
; APPLICANT: Hafner, Stefan
; TITLE OF INVENTION: Method for Zymotic Production of Fine Chemicals Containing
; FILE REFERENCE: Sulphur (Meta)
; CURRENT APPLICATION NUMBER: US/10/525,674
; CURRENT FILING DATE: 2005-02-24
; PRIOR APPLICATION NUMBER: PCT/EP 2003/009452
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: DE 102 39 073.8
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 61
; LENGTH: 5860
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: plasmid
US-10-525-674-61

Query Match 22.0%; Score 1136.4; DB 8; Length 5860;
Best Local Similarity 99.9%; Pred. No. 1.3e-259;
Matches 1137; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3405 GGGCTGCTAAAGGAGCGGAAACAGTAGAAAGCCAGTCCGAGAAACGGTGTGACCCCG 3464
DB 1652 GGGCTGCTAAAGGAGCGGAAACAGTAGAAAGCCAGTCCGAGAAACGGTGTGACCCCG 1711

QY 3465 GATGAATGTACGTACTGGGCTATCTGGACAAAGGAGGAAACGCAAGCGCAAGAGAAAGCA 3524
DB 1712 GATGAATGTACGTACTGGGCTATCTGGACAAAGGAGGAAACGCAAGCGCAAGAGAAAGCA 1771

QY 3525 GGTAGCTTCGATGGGCTTACATGCGGATAGTACGCTGGGCGGTTTATGACAGCAAG 3584
DB 1772 GGTAGCTTCGATGGGCTTACATGCGGATAGTACGCTGGGCGGTTTATGACAGCAAG 1831

QY 3585 CGAACCGGAATGCCAGCTGGGCGCCCTCTGTAAGTTGGGAAGCCCTGCAAGATGAAA 3644

Db 1832 CGAACCGGAATTGCCAGCTGGGGCGCCCTCTGTAAGGTTGGGAAGCCCTGCAGAAATAA 1891
Qy 3645 CTGGATGGCTTCTTGGCCGCAAGGATCTGATGGCGAGGGGATCAAGATCTGATCAAGA 3704
Db 1892 CTGGATGGCTTCTTGGCCGCAAGGATCTGATGGCGAGGGGATCAAGATCTGATCAAGA 1951
Qy 3705 GACAGGATGAGGATGCTTTCGATGATGMAACAAGATGATTTGACGAGGTTCTCCGGC 3764
Db 1952 GACAGATGAGGATGCTTTCGATGATTTGAACAGATGATTTGACGAGGTTCTCCGGC 2011
Qy 3765 CGCTTGGGTGGAGAGGCTATTCCGCTATGACTGGGCACACAGACAAATCGGCTGCTCTGA 3824
Db 2012 CGCTTGGGTGGAGAGGCTATTCCGCTATGACTGGGCACACAGACAAATCGGCTGCTCTGA 2071
Qy 3825 TGCCGCGCTTTCGCGCTGTGACGCGAGGGGCGCCGGTCTTCTTTTGTCAAGACCGACCT 3884
Db 2072 TGCCGCGCTTTCGCGCTGTGACGCGAGGGGCGCCGGTCTTCTTTTGTCAAGACCGACCT 2131
Qy 3885 GTCCGCGTCCCTGAATGAACATGACGAGCGAGCGGCTATCTGCTGGCTGGCCACGAC 3944
Db 2132 GTCCGCGTCCCTGAATGAACATGACGAGCGAGCGGCTATCTGCTGGCTGGCCACGAC 2191
Qy 3945 GGGCGCTTCTTTCGCGAGCTGTCTCGAGCTTGTCTCACTGAAGCGGGAAGGACTGGCTGCT 4004
Db 2192 GGGCGCTTCTTTCGCGAGCTGTCTCGAGCTTGTCTCACTGAAGCGGGAAGGACTGGCTGCT 2251
Qy 4005 ATTGGCGGAAGTGGCGGGGAGGATCTCTCTGTCTATCTCACTTGTCTCTGCGGAGAAAGT 4064
Db 2252 ATTGGCGGAAGTGGCGGGGAGGATCTCTCTGTCTATCTCACTTGTCTCTGCGGAGAAAGT 2311
Qy 4065 ATCCATCATGGCTGATGCAATGGCGGGCTGCATAGCTTGTATCGGCTACTGCCCAAT 4124
Db 2312 ATCCATCATGGCTGATGCAATGGCGGGCTGCATAGCTTGTATCGGCTACTGCCCAAT 2371
Qy 4125 CGACCACCAAGCGAAACATCGCATCGAGCGAGCAGTACTCGGATGGAAGCGGCTTGT 4184
Db 2372 CGACCACCAAGCGAAACATCGCATCGAGCGAGCAGTACTCGGATGGAAGCGGCTTGT 2431
Qy 4185 CGATCAGGATGATCTGGAAGAGATCATGAGGCTCGCGCAGCCGAACTGTTGCGCAG 4244
Db 2432 CGATCAGGATGATCTGGAAGAGATCATGAGGCTCGCGCAGCCGAACTGTTGCGCAG 2491
Qy 4245 GCTCAAGCGCGCATGCCCGAGGAGTCTGCTGTGAGCCCATGCGGATGCTGCTT 4304
Db 2492 GCTCAAGCGCGCATGCCCGAGGAGTCTGCTGTGAGCCCATGCGGATGCTGCTT 2551
Qy 4305 GCGGAATATCATGTTGGAAATGGCGGCTTTTCTGGAATTCATCGACTGTGGCGGCTGG 4364
Db 2552 GCGGAATATCATGTTGGAAATGGCGGCTTTTCTGGAATTCATCGACTGTGGCGGCTGG 2611
Qy 4365 TGTGGCGAGCCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGG 4424
Db 2612 TGTGGCGAGCCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGG 2671
Qy 4425 CGGCGAATGGGCTGACCGCTTCTGCTGTTACGGTATCGCGCTCCCGATTGCGCAGCG 4484
Db 2672 CGGCGAATGGGCTGACCGCTTCTGCTGTTTACGGTATCGCGCTCCCGATTGCGCAGCG 2731
Qy 4485 CATCGCCTTCTATCGCCTTCTTGAGAGTCTTCTTGAGCGGAGCTCTCGGGGTTGATA 4542
Db 2732 CATCGCCTTCTATCGCCTTCTTGAGAGTCTTCTTGAGCGGAGCTCTCGGGGTTGATA 2789

RESULT 15

US-10-525-674-66
; Sequence 66, Application US/10525674
; Publication No. US20060003425A1
; GENERAL INFORMATION:
; APPLICANT: Kroger, Burkhard
; APPLICANT: Zelder, Oskar
; APPLICANT: Kolprogge, Corinna
; APPLICANT: Schroeder, Hartwig

; APPLICANT: Hafner, Stefan
; TITLE OF INVENTION: Method for Zymotic Production of Fine Chemicals Containing
; FILE OF INVENTION: Sulphur (Meta)
; FILE REFERENCE: 13111-00002-US
; CURRENT APPLICATION NUMBER: US/10/525,674
; PRIOR FILING DATE: 2005-02-24
; PRIOR APPLICATION NUMBER: PCT/EP 2003/009452
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: DE 102 39 073.8
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 66
; LENGTH: 5860
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: plasmid
US-10-525-674-66

Query Match 22.0%; Score 1136.4; DB 8; Length 5860;
Best Local Similarity 99.9%; Pred. No. 1.3e-259;
Matches 1137; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 3405 GGGCTGCTAAAGGAAGCGGAACACGTAGAAAGCCAGTCCGAGAAACGCTGCTGACCCCG 3464
Db 1652 GGGCTGCTAAAGGAAGCGGAACACGTAGAAAGCCAGTCCGAGAAACGCTGCTGACCCCG 1711
Qy 3465 GATGAATGTCAAGTACTGGGCTATCTGGACAAAGGAAACCGCAAGCGCAAGAAAGCA 3524
Db 1712 GATGAATGTCAAGTACTGGGCTATCTGGACAAAGGAAACCGCAAGCGCAAGAAAGCA 1771
Qy 3525 GGTAGCTTGCAGTGGGCTTACATGGCGATAGTAGACTGGGCGGTTTTATGGACAGCAAG 3584
Db 1772 GGTAGCTTGCAGTGGGCTTACATGGCGATAGTAGACTGGGCGGTTTTATGGACAGCAAG 1831
Qy 3585 CGAACCGGAATGCGACGCTGGGCGCCCTCTGTTAAGGTTGGGAAGCCCTGCAAGTAAA 3644
Db 1832 CGAACCGGAATGCGACGCTGGGCGCCCTCTGTTAAGGTTGGGAAGCCCTGCAAGTAAA 1891
Qy 3645 CTGATGGCTTCTTTCGCGCAAGGATCTGATGCGCAGGGGATCAAGATCTGATCAAGA 3704
Db 1892 CTGATGGCTTCTTTCGCGCAAGGATCTGATGCGCAGGGGATCAAGATCTGATCAAGA 1951
Qy 3705 GACAGGATGAGGATCGTTTTCGATGATGAAACAGATGGATGGACCGAGGTTCTCCGGC 3764
Db 1952 GACAGGATGAGGATCGTTTTCGATGATGAAACAGATGGATGGACCGAGGTTCTCCGGC 2011
Qy 3765 CGCTTGGGTGGAGAGGCTATTTCGCTATGACTGGGCACAAAGATCGGCTGCTCTGA 3824
Db 2012 CGCTTGGGTGGAGAGGCTATTTCGCTATGACTGGGCACAAAGATCGGCTGCTCTGA 2071
Qy 3825 TGCCGCGCTGTTCCGGCTGTGACGCGAGGGGCGCCGGTCTTCTTTTGTCAAGACCGACCT 3884
Db 2072 TGCCGCGCTGTTCCGGCTGTGACGCGAGGGGCGCCGGTCTTCTTTTGTCAAGACCGACCT 2131
Qy 3885 GTCCGCTGCTTGAATGAACATGACGAGCGAGCGCGGCTATCTGTTGGCTGGCCACGAC 3944
Db 2132 GTCCGCTGCTTGAATGAACATGACGAGCGAGCGCGGCTATCTGTTGGCTGGCCACGAC 2191
Qy 3945 GGGGTTCTTTCGCGAGCTGCTGCTGAGCTTGTACCTGAAAGCGGGAAGGAGCTGGCTGCT 4004
Db 2192 GGGGTTCTTTCGCGAGCTGCTGCTGAGCTTGTACCTGAAAGCGGGAAGGAGCTGGCTGCT 2251
Qy 4005 ATTGGCGGAAGTGGCGGGGAGGATCTCTCTGTCTCATCTCACCTTCTCTGCGGAGAAAGT 4064
Db 2252 ATTGGCGGAAGTGGCGGGGAGGATCTCTCTGTCTCATCTCACCTTCTCTGCGGAGAAAGT 2311
Qy 4065 ATCCATCATGGCTGATGCAATGGCGGCTGCTGATACGCTTGTATCGGCTACTGCCCAAT 4124
Db 2312 ATCCATCATGGCTGATGCAATGGCGGCTGCTGATACGCTTGTATCGGCTACTGCCCAAT 2371
Qy 4125 CGACCACCAAGCGAAACATCGCATCGAGCGAGCAGTACTCGGATGGAAGCGGCTTGT 4184


```
Db 2372 CGACCAACCAAGGGAACATCGATCGAGGAGACAGTACTCGANTGGAGAGCCGGTCTTGT 2431
QY 4185 CGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAG 4244
Db 2432 CGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAG 2491
QY 4245 GCTCAAGGCGGCATGCCCGACGGCGAGGATCTCGTGTGACCCATGGCGATGCCCTGCTT 4304
Db 2492 GCTCAAGGCGGCATGCCCGACGGCGAGGATCTCGTGTGACCCATGGCGATGCCCTGCTT 2551
QY 4305 GCCGAATATCATGTGGAAATGGCGCTTTTCTGATTCATCGACTGTGGCCGCTGGG 4364
Db 2552 GCCGAATATCATGTGGAAATGGCGCTTTTCTGATTCATCGACTGTGGCCGCTGGG 2611
QY 4365 TGTGGCGACCCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGG 4424
Db 2612 TGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGG 2671
QY 4425 CGGCGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCGATTCGACGG 4484
Db 2672 CGGCGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCGATTCGACGG 2731
QY 4485 CATCGCTTCTATCGCTTCTTGGAGAGTTCTTCTGAGCGGACTCTGGGGTTGATA 4542
Db 2732 CATCGCTTCTATCGCTTCTTGGAGAGTTCTTCTGAGCGGACTCTGGGGTTGAAA 2789
```

Search completed: February 27, 2006, 14:15:00
Job time : 1264 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2006, 08:34:41 ; Search time 830 Seconds
(without alignments)
11087.273 Million cell updates/sec

Title: US-10-618-570-1

Perfect score: 5177

Sequence: 1 ttgtaagaccaccaccgta.....agcgggggttttatttgg 5177

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA.*

- 1: /cgn2_6/ptodata/1/ina/1/COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5/COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A/COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B/COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/H/COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/PP/COMB.seq.*
- 7: /cgn2_6/ptodata/1/ina/RE/COMB.seq.*
- 8: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2078.8	40.2	7699	3 US-09-645-004-1	Sequence 1, Appli
2	1661.6	32.1	8316	2 US-07-753-520B-4	Sequence 4, Appli
3	1629.4	31.5	5594	3 US-09-380-190A-29	Sequence 29, Appli
4	1588.6	30.7	5130	3 US-09-897-511A-9	Sequence 9, Appli
5	1486	28.7	6365	2 US-08-352-990-1	Sequence 1, Appli
6	1486	28.7	6620	3 US-08-786-531B-3	Sequence 3, Appli
7	1486	28.7	9115	2 US-07-753-520B-3	Sequence 3, Appli
8	1455.2	28.1	5865	3 US-08-654-737B-3	Sequence 3, Appli
9	1448.4	28.0	8591	2 US-08-462-859A-6	Sequence 6, Appli
10	1448.4	28.0	8591	2 US-08-462-859A-8	Sequence 8, Appli
11	1448.4	28.0	8591	2 US-08-123-659A-6	Sequence 6, Appli
12	1448.4	28.0	8591	2 US-08-123-659A-8	Sequence 8, Appli
13	1448.4	28.0	8591	2 US-08-464-247A-6	Sequence 6, Appli
14	1448.4	28.0	8591	2 US-08-464-247A-8	Sequence 8, Appli
15	1448.4	28.0	8591	2 US-08-464-248A-6	Sequence 6, Appli
16	1448.4	28.0	8591	2 US-08-464-248A-8	Sequence 8, Appli
17	1434	27.7	6700	3 US-09-759-152A-1	Sequence 1, Appli
18	1434	27.7	8518	3 US-09-654-449-3	Sequence 3, Appli
19	1434	27.7	8518	3 US-09-759-152A-4	Sequence 4, Appli
20	1434	27.7	8518	3 US-09-654-499-3	Sequence 3, Appli
21	1429.2	27.6	6700	3 US-09-654-449-1	Sequence 1, Appli
22	1429.2	27.6	6700	3 US-09-654-499-1	Sequence 1, Appli
23	1425	27.5	5689	3 US-09-508-516-1	Sequence 1, Appli
24	1171.8	22.6	3176	9 5212080-1	Patent No. 5212080

ALIGNMENTS

RESULT 1

US-09-645-004-1
; Sequence 1, Application US/09645004
; Patent No. 6608238
; GENERAL INFORMATION:
; APPLICANT: Gavora, Jan S.
; APPLICANT: Falconer, Marcia M.
; APPLICANT: Nguyen, Thuy H.
; APPLICANT: Benkel, Bernhard F.
; FILE OF INVENTION: Trans-Somatics With Gene Transfer Into Mammary Epithelial Cells
; TITLE REFERENCE: GALA-06402
; CURRENT APPLICATION NUMBER: US/09/645.004
; CURRENT FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Plasmid pPLAIN
US-09-645-004-1

Query Match	40.2%	Score 2078.8;	DB 3;	Length 7699;
Best Local Similarity	64.3%	Pred. No. 0;		
Matches 3339;	Conservative	0;	Mismatches 1814;	Indels 38; Gaps 14;
Qy	1	TTTGAAGAGCCCCACCGTAGGTGGCAAGCTTAAGTAAACCCACTTTTGCAGGCAT	60	
Db	175	TTTGAAGAGCCCCACCGTAGGTGGCAAGCTTAAGTAAACCCACTTTTGCAGGCAT	234	
Qy	61	GGAAATATACATACTGAGATAGAAAAGTTCAGATCAAGTCAGGCAAGAAACAGC	120	
Db	235	GGAAATATACATACTGAGATAGAAAAGTTCAGATCAAGTCAGGCAAGAAACAGC	294	
Qy	121	TGAATACCAACAGGATATCTGTGTAAGCGGTTCCTGCCCGCTCAGGGCCCAAGAAC	180	
Db	295	TGAATACCAACAGGATATCTGTGTAAGCGGTTCCTGCCCGCTCAGGGCCCAAGAAC	354	
Qy	181	GATGAGACAGCTGAGTATCGGCAACACAGATATCTGTGTAAGCGGTTCCTGCCCGG	240	
Db	355	GATGAGACAGCTGAGTATCGGCAACACAGATATCTGTGTAAGCGGTTCCTGCCCGG	414	
Qy	241	CTCGGGCCCAAGAACAGATGTCCTCCAGATCGGTCAGGCCTCAGCAGTTTCTAGTGA	300	
Db	415	CTCGGGCCCAAGAACAGATGTCCTCCAGATCGGTCAGGCCTCAGCAGTTTCTAGTGA	474	

QY 301 TCATCAGATGTTTCCAGGGTCCCAAGGACCTGAAATGACCCCTGTACCTTATTTGAAC 360
DB |||||
QY 475 TCATCAGATGTTTCCAGGGTCCCAAGGACCTGAAATGACCCCTGTACCTTATTTGAAC 534
DB |||||
QY 361 TAAACAAATCAGTTCGCTTCTCGCTTCTGTTGCGCGCTTCCGCTCTCCGAGCTCAATAAA 420
DB |||||
QY 535 TAAACAAATCAGTTCGCTTCTCGCTTCTGTTGCGCGCTTCCGCTCTCCGAGCTCAATAAA 594
DB |||||
QY 421 AGAGCCCAACCCCTCACTCGGCGCGCCAGTTCCTCGATGACCTGCGTCCCGGGTAC 480
DB |||||
QY 595 AGAGCCCAACCCCTCACTCGGCGCGCCAGTTCCTCGATGACCTGCGTCCCGGGTAC 654
DB |||||
QY 481 CCGTATTCCTCAATAAAGCCTCTTGCTGTTTGGATCCGAATCGTGTCTCGCTGTTCCCTTG 540
DB |||||
QY 655 CCGTATTCCTCAATAAAGCCTCTTGCTGTTTGGATCCGAATCGTGTCTCGCTGTTCCCTTG 714
DB |||||
QY 541 GGAGGGTCTCTCTGAGTGAATGACTACCCAGCAGGGGGCTTTTCATTTTGGGGCTCGT 600
DB |||||
QY 715 GGAGGGTCTCTCTGAGTGAATGACTACCCAGCAGGGGGCTTTTCATTTTGGGGCTCGT 774
DB |||||
QY 601 CCGGGAATTTGGAGACCCCTGCCAGGGACCCAGACCCACCCAGGGAGTAACTGGCC 660
DB |||||
QY 775 CCGGGAATTTGGAGACCCCTGCCAGGGACCCAGACCCACCCAGGGAGTAACTGGCC 834
DB |||||
QY 661 AGCAACTTATCTGTCTGTCCGATTTGTCTAGTGTCTATGTTGATGTTATGCGCCTGCG 720
DB |||||
QY 835 AGCAACTTATCTGTCTGTCCGATTTGTCTAGTGTCTATGTTGATGTTATGCGCCTGCG 894
DB |||||
QY 721 TCTGTACTAGTTAGTAACTAGTCTGTATCTGCGGACCCGTTGGTGAACCTGACGAGTT 780
DB |||||
QY 895 TCTGTACTAGTTAGTAACTAGTCTGTATCTGCGGACCCGTTGGTGAACCTGACGAGTT 954
DB |||||
QY 781 CTGAAACCCCGCCCAACCTTGGGAGAGCTCCAGGAGACTTTGGGGCCGTTTTTGTGG 840
DB |||||
QY 955 CTGAAACCCCGCCCAACCTTGGGAGAGCTCCAGGAGACTTTGGGGCCGTTTTTGTGG 1014
DB |||||
QY 841 CCGGACCTTGAGGAAGGAGTGCATGTGGAATCCGACCCCGCTCAGGATATGTGTTCTGTT 900
DB |||||
QY 1015 CCGGACCTTGAGGAAGGAGTGCATGTGGAATCCGACCCCGCTCAGGATATGTGTTCTGTT 1074
DB |||||
QY 901 AGGAGACGAGAACCTTAAACAGTTCCTCCGCTCCGTTCTGAAATTTTGTCTTCGTTTGGAA 960
DB |||||
QY 1075 AGGAGACGAGAACCTTAAACAGTTCCTCCGCTCCGTTCTGAAATTTTGTCTTCGTTTGGAA 1134
DB |||||
QY 961 CCGAAGCCCGCTTCTGTCTGCTCAGCAGAGCTTGGGCTGCAGCTCGACTCTAGAGGA 1020
DB |||||
QY 1135 CCGAAGCCCGCTTCTGTCTGCTCAGCAGCTGTCAGAGCTGTTCTGTCTGTCTGTCT 1194
DB |||||
QY 1021 TCAATTCGGCAGAGTAAATCGGTCTGCTGCTCTTTAGGACATATGAAGTATGGCAGGT 1080
DB |||||
QY 1195 GACTGTGTTTCTGTATTTGTCTGAAATTTAGGGCCAGACTGTTACCACTCCCTTAAGTTT 1254
DB |||||
QY 1081 GGGATGACTTTCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTTCG 1140
DB |||||
QY 1255 GACCTTAGTCTACTGGAAGAATGTCGAGCGGATCGCTCACACCACTCGGTAGATGTCAA 1314
DB |||||
QY 1141 ATGCTCGCTCAGTGACAGCTTGTCTCTCCCATCTTAAATGGCCCACTCTTCAAG 1200
DB |||||
QY 1315 GAAAGAGCGTTGGGTATCTTCTGTCTGTCGAGAATGGCCACCTTTAACTGCGATGGCC 1374
DB |||||
QY 1201 AGAGGATGAAGTCTTATAAATCTGACCTGATCACCCTTTTATCTCATTTGTGTTGTAGTTC 1260
DB |||||
QY 1375 GCGAGACGGCACCTT-TAAACGAGACCTCTATCACCAGGTTAAGATCAAGTCTT-TTC 1431
DB |||||
QY 1261 TCGTGCCCATCATTTGGCATAGTGGCAGCTCAGCTCTCGAATGGGAAACGAAAGTATGCA 1320
DB |||||
QY 1432 ACCTGGCCCGCATGGACACCCAGACCAGTCCCTTACATCTGTGACTTGGGAAGCTTGGC 1491
DB |||||
QY 1321 CCGTGTGGCTCAGTTAATGAGATATATCTCCAGTCCGGAAGGCAAGGAATGGCAGTG 1380
DB |||||
QY 1492 TTTTGACCCCTCTCCCTGGGTCAAGCCCT-----TTGTACACCTTAAGCCCTCCGCTCCT 1546
DB |||||
QY 1381 AAGATGAATGAGATTTTCGAGAAGCTGTGATGGAAAGCATGAGCAACATGGAAGACGAA 1440

DB |||||
QY 1547 CTTCTCCATCCGCCCCGCTCTCTCCCTTGAAACCTCTCTGTTGACCCCGCTCGATCC 1606
DB |||||
QY 1441 TCCAGTATCTTTTTCAGATAATGAAGCCAAATCTCTCTAGATGCTGAAGAAATTTCCAAATTTCA 1500
DB |||||
QY 1607 TCCCTTTATCCAGCCCTCACTCTCTCTTAGGCCCGGAAATTCGTTAACTCGACATGAA 1666
DB |||||
QY 1501 GCATAACAACTGATCAAGATTTAAATGATGTTCTTTTCCAGCTAAATTCCTTACTTCTCCT 1560
DB |||||
QY 1667 GTCCTTCTCTCTCTCGCAGCTGTGCGGCTTTTGGGCAAGTACAATCCCAACTCAG 1726
DB |||||
QY 1561 CCATCCAGG--AACATGAGAAATATCATAGGGGATATCTCCAAGTCAATTAGTAGGTCTGAA 1618
DB |||||
QY 1727 GCTGGGAGGACATCTATCTGTCATCTCTTTGAAATGGCGCTGGGCCGACATGTCACCTGAG 1786
DB |||||
QY 1619 CACCACAGTACTTGAATTTGACAGTTTCACTATTAATGAACTGTAATGGCAGAGTCCCAAGAGAA 1678
DB |||||
QY 1787 TGCGAACACTATTTAGCTCTCTAATGGGTTTGGAGAGTTTCAAGGTTTCTCTCTCCAAATGAA 1846
DB |||||
QY 1679 TGCATTTTAAACAAACAGAGGAGATGCGTAAATTTAGAGGAGCGTATATACAAATGATCAGC 1738
DB |||||
QY 1847 AACATGTCATTAATCTCGAACAGGCCCTGTTGGGAAAGATACCAAGCCCATCAGCTAC 1906
DB |||||
QY 1739 AGAAATTAAGTCTCTAGATGAAACAAAGTATATTTTGGAAACAGGAAATTAAGGGGAAAT 1798
DB |||||
QY 1907 AAGATCTGCAGTCGATCGGGCAATGAAATGAATTTTCAGAGACATGTTGACCCAGATGCAAC 1966
DB |||||
QY 1799 GAAACTGTTGAATAATTAATCAATGATCTGAGGCTGAAAGGATTTGGGAAACATTTCTCAGAC 1858
DB |||||
QY 1967 AATGTTGGAG--TTGCTATTTATGTTGATGCTGTTGTTCAATCACAATGTTGATCTATGGG 2025
DB |||||
QY 1859 ATTGAAAAATATCACTTTTACTCCAAGTGGCAGAAAGTGTCTCGCTGACTGGGAAATGGAC 1918
DB |||||
QY 2026 TGGCACGGGACCCCACTCAACATGTGGGAGCTATTTCAACACCCGGGACTAGAGATTTTCC 2085
DB |||||
QY 1919 CAACGATCTGGGCTCCCAACATGACCATCGGGGCTGTGAAACAGCAGAGAGTGAAATTCACAGG 1978
DB |||||
QY 2086 CGCTGTGCGGTACTCTGCTGGGATTTCAATGACGGCAATGTCACTGCTCAAGTGGAGA 2145
DB |||||
QY 1979 CACTTACATCAGAGCCGTAAACAGCCACATCAATATGAGATCAAGAGTCCACACTGCTATGG 2038
DB |||||
QY 2146 CATCGAAAAATTTATGGGACATGTATCAGGTCCGGGATTTGCAAGTTGTCCAGCCTTCTTGA 2205
DB |||||
QY 2039 GACA--CAAAACACCATCAACAGAGGACCCAGCCCACTTTGGCTTCACTGCTCAATTTGGA 2097
DB |||||
QY 2206 TCTGGCTCTGGAGAAGACTATGTACGCTCAACAAATTTGACGCTACATGAATACCTCAT 2265
DB |||||
QY 2098 AGTTTTTTCAGAGTCCACCACTGTCTTTCAGGGGCCAGTGTCTTATAGACAGGAAATGGGAAGG 2157
DB |||||
QY 2266 TGATATGGGTGAGCAGGGTTCCGGATCGATGCTGCCAAGCATATGTGGCCAGGGACAT 2325
DB |||||
QY 2158 AGGTCTGAAAGACCATGTGGTGTCTGGGTCAAGTGTATGATGATTTGGTGTATGACTGGA 2217
DB |||||
QY 2326 AAGAGCGTTTCTGAGCAAACTGCACGATCTAAATACTCAGTGGTGTTCAGCAGGAAACGAA 2385
DB |||||
QY 2218 AAGCTACAGGGTC--GCCATCAACATCTTCACTCGCTCGGCAACAGAGAGTGTAGTG 2276
DB |||||
QY 2386 ACCCTTTATTTACCAAGGTAATTTGACTTTGGGAGGAGAGCCAAATCAGGCACTCAGTA 2445
DB |||||
QY 2277 AGTGACCAAGGCTCTCTG--GACTCCAGGTGAAAGAGGAGATAGAGGCCCTCTCTGGAC 2333
DB |||||
QY 2446 CTTTGGGAATGGCGCTGACAGNAATTCAGATGTTGGTCCAACTGCGGACGGTGTATCCG 2505
DB |||||
QY 2334 AAAATGTTATACAGGCTTTCCAGGTCTATAGGTACTCCAGGTCTTAAAGGTGATCGGG 2393
DB |||||
QY 2506 GAAGTGAATGGAGAGAAGATGGCTTACTTTAAAGAACTGGGGAGAAAGCTGGGGCTTTGT 2565
DB |||||
QY 2394 GGGATCTCTGTTTACCTGGAGTTCCAGGATTTCCAGGAC--CAATGGGGAAGACCCGGA 2451
DB |||||
QY 2566 GCCTTCTGACAGAGCCCTCTGGTGTGTTGTTGATTAACACGACAAACCGCGGGGGACGGGGC 2625
DB |||||
QY 2452 AGCCAGGACTTAATTTGGAACAAAAGGCCAGAGGGAGAAAAAGGAGGTGGAGCATGCAAA 2511
DB |||||

Db 2626 AGCGGAGCTCCATTCTTCTTCTGGGATGCCAGGCTTTTATAAATGGCGGTGGTTT 2685
Qy 2512 GACAACTTAATACAGTCCGATCTGTGGTGGGAGCGGCCCTTCA CGAAGGACAGATGGAGA 2571
Db 2686 CATGCTCGCTCATCGTACGGGTTTACACACGGGTGATGTCAAGATTATCGTTGGCCAAAGATA 2745
Qy 2572 TTTTTCAGAGGCGCAGTGGGTACGGTGTGTGACGACCGCTGGGAACTCGGTGGAGAC 2631
Db 2746 TTTTTCAGAGGCGCAGTGGGTACGGTGTGTGACGACCGCTGGGAACTCGGTGGAGAC 2805
Qy 2632 T-----GGTGTGTGAGGAGCTTGGGATACAAAGGTGTCAAAGTGTGCATAAGCAGCT 2687
Db 2806 GACGAAGTCCGTATCAATCAAGCAGACACTACCTGTGGCAATGACTGGGTCTGGGAACA 2855
Qy 2688 TATTTTGAAGAGTACGGGTCCAATATGGCTGAATGAAGTATTTTGTTCGGGAAGAG 2747
Db 2866 TCGCTGGCGCAAAATAAG-----AACATGGTTATCTCCGTAAATGTGTAGACGTCAAG 2921
Qy 2748 TCATCCATTGAAGAGTGCAGAAATAGACAGTGGGGTGTGAGAGCTGTTCACACGAGAA 2807
Db 2922 CTTTCTCAAACTGGTGGGACAAACGGGAGCAATCAAGTAGCTTTTCGGTCCGGCGACAGAG 2981
Qy 2808 GATGCTGGGGGTCACTTTGACCTACATAATATGATCATATTTTCAATCACAATTTTTTAAA 2867
Db 2982 GCTTCATTTGTCTTAATAATGATGACTGGTATATGAATGTGCAATTTGCAAACTGGTCTGC 3041
Qy 2868 CTGTATTAAGTGAATTTTTTCTTGTCTTCAATAAATCAGCTTAATTAATATTTAAGA 2927
Db 3042 CTGTGGAACTTACTCGCATGTATTTCTGACAAAAGGAGGAGTGCGTGTACTGAA 3101
Qy 2928 AACTAAGAAATTTATCCACAGAAAGGAATATTTAAATACTCACTGGATAACATATAAAA 2987
Db 3102 AGCAGGTGTACGTCTTCCTCGATGGAAAGGCCAATTTCCAGATTAGTAACAGCGATGAAG 3161
Qy 2988 TAGCTTCATATTTGCTTCAAAATACAGAAACCATTTCAACTTCTCTAGTGTTTTAAAGTGC 3047
Db 3162 ATCAATTTGTGCAATTCAGTGTATGATGCAAGTTATAGTTCGAGGATCCACTAGTAAC 3221
Qy 3048 TCGTGCCGAATTTGATCCCTCAGGATATAGTATTTTCTGCTTTTGCATAGGAGGGGAAA 3107
Db 3222 GGCGCCAGTGTGTGGAATTCGG-----CTTGTGACATCTAGGGCGGCCAATTCGCC 3276
Qy 3108 TGTAGTCTTATGCAATCTCTTGTAGTCTTGCAACATGTTAAGATGAGTTAGCAACATG 3167
Db 3277 CCTCTCCCGCCCGCCCTTAAGCTTACTGGCGAAGCGCTTGGAAATAAGGCGCGGTGTG 3336
Qy 3168 CTTTACAGGAGAGAAAAGCACCGTGCATGCCGATTTGGTGAAGTAAGTGTACGATC 3227
Db 3337 TTTGTCTATATGTGATTTTCCACCATATGTCCGCTTTTGGCAA-----TGTGAGGGCC 3390
Qy 3228 GTGCTTTATAGGAAGGCAACAGACGGGTCTGACATGATTGGAGCAACCACTGAATTC 3287
Db 3391 CGAAACCTGGCCCTGTCTTCTTGNCGAGCATTCCTAGGGGTCTTTCCCTCTCGCCAAA 3450
Qy 3288 GCATTCAGAGATATTTGATTTAAGTGTCTAGTCTGATACAGCAAAACGCCATTTGACCAT 3347
Db 3451 GGAATGCAAGGTCTGTGAAATGTGTGAAAGAGCAGTCTCTCTGGAAGCTTCTTGAAGA 3510
Qy 3348 TCACCATATTTGGTGTGCACCTTCAAGCTTACGCTGCCGACGACTCAGGGCGCAAGGG 3407
Db 3511 CAAACAAGCTGTAGCGACCTTTTGCAGGCGAGCGGAACCCCCACCTGGCGACAGGTGC 3570
Qy 3408 CTGCTAAAGGAAGCGGAACAGTGAAGAGCCAGTCCGACAGAAACGGGTCTGACCCCCGAT 3467
Db 3571 CTCTGGGCCAAAGCCAGGTGTATAGATACACTGCAAGGGCGGCAACCCACAGTGC 3630
Qy 3468 GAATGTACGTACTGGGTCTATCTGACA-AGGGAAGCGCAAGCGCAAGAGAACAGG 3526
Db 3631 CACGTTGTGAGTGTGATAGTTGTGNAAGAGTCAAAATGGCTCTCTCTCAAGGTAGTCAAC 3690
Qy 3527 TAGCTTGCAGTGGGTCTTACATGGCGATAGTAGTGGCGGCTTTTATGGACAGCAAGCG 3586
Db 3691 AAGGGGCTGAAGGATGCGCCAGAGGTACCCCAATTTGTATGGGAATCTGATCTGGGGCCCTCG 3750

Qy 3587 AACCGGAATTCAGCTGGGGCGCCCTCTCTGTAGAGTTGGGAAGCCCTTGCAAAAGTAAACT 3646
Db 3751 GTGCACATGCTTTACATGTGTGTAGTCAGGTTAAAAAGCTCTAGGCCCCCAGAACAC 3810
Qy 3647 GGATGGCTTTCTTGGCCCGCAGGATCTGATGGCCGAGGGATCAAGATCTGATCAAGAGA 3706
Db 3811 GGGGACGTGGTTTCTTTGAAAAACACAGATGATAGCTTGGCCACAACCCGGGATATTC 3870
Qy 3707 CAGGATGAGGATCGTTTTCGATGATTAACAAGATGATTCACGAGGTTCTCCGGCCG 3766
Db 3871 CTGCAGCAATATGGGATCGGCATTCGAACAGATGATTCGACGCGAGGTTCTCCGGCCG 3930
Qy 3767 CTTGGGTGGAGAGCTATTTGGCTATGACTGGGCAACAAGCAATCGGTGCTCTGATG 3826
Db 3931 CTTGGGTGGAGAGCTATTTGGCTATGACTGGGCAACAAGCAATCGGTGCTCTGATG 3990
Qy 3827 CCGCGGTGTTCCGGCTGTGACGAGGGGCGCGGTTCTTTTGTCAAGACCGACCTGT 3886
Db 3991 CCGCGGTGTTCCGGCTGTGACGAGGGGCGCGGTTCTTTTGTCAAGACCGACCTGT 4050
Qy 3887 CCGGTGCTGAAATGAATCTGACGAGCGAGCGCGGCTATCGTGGCTGGCCACGACGG 3946
Db 4051 CCGGTGCTGAAATGAATCTGACGAGCGAGCGCGGCTATCGTGGCTGGCCACGACGG 4110
Qy 3947 GCGTTCCTTGGCAGCTGTCTGACGCTTGTCACTGAAGCGGGAAGGACTTGGCTGTAT 4006
Db 4111 GCGTTCCTTGGCAGCTGTCTGACGCTTGTCACTGAAGCGGGAAGGACTTGGCTGTAT 4170
Qy 4007 TGGCGAAGTGGCGGGCAGGATCTCTGTCTCATCTCACTTGTCTCTGCGGAGAAAGTAT 4066
Db 4171 TGGCGAAGTGGCGGGCAGGATCTCTGTCTCATCTCACTTGTCTCTGCGGAGAAAGTAT 4230
Qy 4067 CCATCATGGCTGATGCAATCGCGGCTGATAGCTTGTGATCCGGCTACCTGCCCATTCG 4126
Db 4231 CCATCATGGCTGATGCAATCGCGGCTGATAGCTTGTGATCCGGCTACCTGCCCATTCG 4290
Qy 4127 ACCACCAAGGCAAACTCGCATCGAGCAGCAGCTACTCGGATGGAAGCCGGTCTTGTG 4186
Db 4291 ACCACCAAGGCAAACTCGCATCGAGCAGCAGCTACTCGGATGGAAGCCGGTCTTGTG 4350
Qy 4187 ATCAGGATGATTCGGAGCAAGAGCATCAGGGGCTCGCGGCGCAAGCACTGTTCCGACGC 4246
Db 4351 ATCAGGATGATTCGGAGCAAGAGCATCAGGGGCTCGCGGCGCAAGCACTGTTCCGACGC 4410
Qy 4247 TCAAGGCGCATGCGCGGAGGATCTCTGCTGTGACCCATGCGGATGSCCTGCTTGC 4306
Db 4411 TCAAGGCGCATGCGCGGAGGATCTCTGCTGTGACCCATGCGGATGSCCTGCTTGC 4470
Qy 4307 CGAATATCATGGTGGAAAAATGGCGCTTTTCTGGATTTCATCGACTGTGGCGCGTGGGTG 4366
Db 4471 CGAATATCATGGTGGAAAAATGGCGCTTTTCTGGATTTCATCGACTGTGGCGCGTGGGTG 4530
Qy 4367 TGGCGGACCGCTATCAGGACATAGCGTTGGGTACCCGTGATATTCGTGAAGAGCTTGGCG 4426
Db 4531 TGGCGGACCGCTATCAGGACATAGCGTTGGGTACCCGTGATATTCGTGAAGAGCTTGGCG 4590
Qy 4427 GCGAATGGGCTGACCGCTCTCTGCTGCTTTCAGGATTCGCGCTCCGATTCGAGCGCA 4486
Db 4591 GCGAATGGGCTGACCGCTCTCTGCTGCTTTCAGGATTCGCGCTCCGATTCGAGCGCA 4650
Qy 4487 TCGCTTTCTATCGCTTTCTTGAAGAGTTCCTTCTGAGCGGAGCTCTGGGGTTCGATAAAT 4546
Db 4651 TCGCTTTCTATCGCTTTCTTGAAGAGTTCCTTCTGAGCGGAGCTCTGGGGTTCGATAAAT 4710
Qy 4547 AAAAGATTTTATTTAGTCTCCAGAAAAGGGGGAATGAAAGACCCCACTGTAGGTTG 4606
Db 4711 AAAAGATTTTATTTAGTCTCCAGAAAAGGGGGAATGAAAGACCCCACTGTAGGTTG 4770
Qy 4607 GCAAGCTAGCTTAAGTAACGCCATTTTGAAGCATCGGAATAATACATACTGAGATAG 4666
Db 4771 GCAAGCTAGCTTAAGTAACGCCATTTTGAAGCATCGGAATAATACATACTGAGATAG 4830

```
Qy 4667 AGAAGTTTCAGATCAAGGTGAGGAGCAGATGAGGAGCAGCTGAATATATGGCCAAACAGGATAT 4726
Db 4831 AGAAGTTTCAGATCAAGGTGAGGAGCAGATGAGGAGCAGCTGAATATATGGCCAAACAGGATAT 4890
Qy 4727 CTGTGTTAAGCAGTTCTGTCGCCCGGCTCAGGGCCAGAACAGATGGAACAGCTGAATATG 4786
Db 4891 CTGTGTTAAGCAGTTCTGTCGCCCGGCTCAGGGCCAGAACAGATGGAACAGCTGAATATG 4950
Qy 4787 GGCACAAACAGGATATCTGTGTGTAAGCAGTTCTGTCGCCCGGCTCAGGGCCAGAACAGATG 4846
Db 4951 GGCACAAACAGGATATCTGTGTGTAAGCAGTTCTGTCGCCCGGCTCAGGGCCAGAACAGATG 5010
Qy 4847 GTCCCCAGATCGGGTCAGACCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGT 4906
Db 5011 GTCCCCAGATCGGGTCAGACCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGT 5070
Qy 4907 GCCCAAGGACCTGAAATGACCCCTGTGCTTATTTGAACCTAACCAATCAGTTTCCTCTC 4966
Db 5071 GCCCAAGGACCTGAAATGACCCCTGTGCTTATTTGAACCTAACCAATCAGTTTCCTCTC 5130
Qy 4967 GCTTCTGTTCGGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTC 5026
Db 5131 GCTTCTGTTCGGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTC 5190
Qy 5027 GGGCGCCAGTCTCCGATTGACTGAGTCGCGCGGTACCCGTGTATCCAAATAAACCCCTC 5086
Db 5191 GGGCGCCAGTCTCCGATTGACTGAGTCGCGCGGTACCCGTGTATCCAAATAAACCCCTC 5250
Qy 5087 TTGCAGTTGCAATCCGACTGTGTGGTCTCGCTGTTCTTGGAGGCTCTCCTCTGAGTGATT 5146
Db 5251 TTGCAGTTGCAATCCGACTGTGTGGTCTCGCTGTTCTTGGAGGCTCTCCTCTGAGTGATT 5310
Qy 5147 GACTACCCGTGAGCGGGGTCTTTCATTTGG 5177
Db 5311 GACTACCCGTGAGCGGGGTCTTTCATTTGG 5341
```

RESULT 2

```
US-07-753-520B-4
; Sequence 4, Application US/07753520B
; Patent No. 5352595
; GENERAL INFORMATION:
; APPLICANT: Tapscott, J.; Weintraub, H.M.; Palmer, T.D.
; TITLE OF INVENTION: "MYOD REGULATORY REGION"
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
; STREET: 2800 Pacific First Center, 1420 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette-5.25 inch, 1.2Mb storage
; COMPUTER: IBM PC/386 Compatible
; OPERATING SYSTEM: MS-DOS 4.01
; SOFTWARE: Word for Windows-t
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07753,520B
; FILING DATE: 19910903
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: none
; FILING DATE: none
; ATTORNEY/AGENT INFORMATION:
; NAME: Sundemo, John, S.
; REGISTRATION NUMBER: 34,446
; REFERENCE/DOCKET NUMBER: FHCR-1-5789
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 1-206-682-8100; 1-206-224-0727 (direct)
; TELEFAX: 1-206-224-0779
; TELEX: 4938023
; INFORMATION FOR SEQ ID NO: 4:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8316 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other; plasmid DNA
; DESCRIPTION: PLHDMN-NSA; 5'LTR (position 1-1159); y+ (position 1160-1640); HisD (position 1641-2928); Myo-D NSA Apal fragment (position 2929-4389); driving neo (position 4390-5259); with 3'LTR (position 5260-5964); Figures 8A-8C.
; US-07-753-520B-4

Query Match 32.1%; Score 1661.6; DB 2; Length 8316;
Best Local Similarity 61.1%; Pred. No. 0;
Matches 3536; Conservative 0; Mismatches 1641; Indels 607; Gaps 29;

Qy 1 TTTGAAAGACCCACCCGTAGGTGGCAAGCTAGCTTAAAGTAAACGCCACTTTTGCAGGCAAT 60
Db 175 TTTGAAAGACCCACCCGTAGGTGGCAAGCTAGCTTAAAGTAAACGCCACTTTTGCAGGCAAT 234
Qy 61 GGAATAATACATAAATCAGATAGAAAGTTTCAATCAAGTCAAGGTCAAGAACAAAGAAACAGC 120
Db 235 GGAATAATACATAAATCAGATAGAAAGTTTCAATCAAGTCAAGGTCAAGAACAAAGAAACAGC 294
Qy 121 TGAATACAAACAGGATATCTGTGTAAAGCGGTTCTGCCCCCGGCTCAGGGCCCAAGAAACA 180
Db 295 TGAATACAAACAGGATATCTGTGTAAAGCGGTTCTGCCCCCGGCTCAGGGCCCAAGAAACA 354
Qy 181 GATGAGACAGCTGAGTGATGGGCAAAACAGGATATCTGTGTAAAGCGGTTCTGCCCCCGG 240
Db 355 GATGAGACAGCTGAGTGATGGGCAAAACAGGATATCTGTGTAAAGCGGTTCTGCCCCCGG 414
Qy 241 CTGGGGCCAAAGAACAGATGTCCTCCAGATGCGGTCAGCCCTCAGCAGTTTCTAGTGAA 300
Db 415 CTGGGGCCAAAGAACAGATGTCCTCCAGATGCGGTCAGCCCTCAGCAGTTTCTAGTGAA 474
Qy 301 TCATCAGATGTTTCCAGGGTGCCTCCAGAGACCTGAAATGACCCCTGTACCTTATTTGAAC 360
Db 475 TCATCAGATGTTTCCAGGGTGCCTCCAGAGACCTGAAATGACCCCTGTACCTTATTTGAAC 534
Qy 361 TAACCAATCAGTTGCTGCTTCTGCTTCTGTTGCGCGGCTTCCGCTCTCCGAGCTCAATAAA 420
Db 535 TAACCAATCAGTTGCTGCTTCTGCTTCTGTTGCGCGGCTTCCGCTCTCCGAGCTCAATAAA 594
Qy 421 AGAGCCCAACCCCTCAGCTCGGCGGCGCAGTCTCCGATAGACTGCTGCCCGGCTAC 480
Db 595 AGAGCCCAACCCCTCAGCTCGGCGGCGCAGTCTCCGATAGACTGCTGCCCGGCTAC 654
Qy 481 CCGTATTTCCCAATAAAGCCCTTCTGCTGTTTGCATCCGAATCGTGGTCTCGCTGTCTTGTG 540
Db 655 CCGTATTTCCCAATAAAGCCCTTCTGCTGTTTGCATCCGAATCGTGGTCTCGCTGTCTTGTG 714
Qy 541 GAGAGGCTCTCTGAGTGATGATGATACCAAGAGCGGGGTCTTTTCAATTTGGGGGCTCGT 600
Db 715 GAGAGGCTCTCTCTGAGTGATGATGATACCAAGAGCGGGGTCTTTTCAATTTGGGGGCTCGT 774
Qy 601 CCGGATTTGGAGACCCCTCGCCAGGACCAACCGACCCACCGGAGGTAACTCGGCC 660
Db 775 CCGGATTTGGAGACCCCTCGCCAGGACCAACCGACCCACCGGAGGTAACTCGGCC 834
Qy 661 AGCAACTTATCTGTGCTGTCTGATGCTAGTGTCTATGTTTGTATGATGCTATGCGCCTGCG 720
Db 835 AGCAACTTATCTGTGCTGTCTGATGCTAGTGTCTATGTTTGTATGATGCTATGCGCCTGCG 894
Qy 721 TCTGTACTAGTAACTAGCTTAACTAGCTCTGTATCTGCGGACCCGCTGGTGAACGAGTT 780
Db 895 TCTGTACTAGTAACTAGCTTAACTAGCTCTGTATCTGCGGACCCGCTGGTGAACGAGTT 954
Qy 781 CTGAACACCCGCGCAACCCCTCGGAGACCTCCAGGAGCTTTGGGGGCGCTTTTGTGG 840
Db 955 CTGAACACCCGCGCAACCCCTCGGAGACCTCCAGGAGCTTTGGGGGCGCTTTTGTGG 1014
Qy 841 CCGACCTGAGGAAGGAGTCGATGTGGAATCCGACCCCGTCAAGATATGTGTTCTGGT 900
```

Db 1015 CCGGACCTGAGAGGAGTGCATGTGAATCCGACCCCGTCAGGATATGGTTCTGGT 1074
Qy 901 AGGAGACGAGAACCTTAACAGAGTCTCCCGCTCCGCTGAAATTTTGTCTTTCGGTTGGAA 960
Db 1075 AGGAGACGAGAACCTTAACAGAGTCTCCCGCTCCGCTGAAATTTTGTCTTTCGGTTGGAA 1134
Qy 961 CCGAAGCCGCGGCTTCTGCTGCTGAGCAGCAAGCTTTGGGCTGAGAGTGCATCTAGAGGA 1020
Db 1135 CCGAAGCCGCGGCTTCTGCTGCTGAGCAGCAAGCTTTGGGCTGAGAGTGCATCTAGAGGA 1134
Qy 1021 TCAATTCGGCAGAGTAATCGGTGCTGCTGCTTTAGGACATATGAAGTATGSC----- 1075
Db 1195 TTCTGTATTTCTGAAAATTAGGCCAGACTGTACCACTCCCTTAAGTTTACCTTAG 1254
Qy 1076 -----ACAGTGGGATGACTTTCTGATCAGCAGAGAGACACTGACAGCTGTACAG 1125
Db 1255 GTCACCTGGAAAGATCTCGAGCGGATCGCTCACACCACTCGGTAGATGTCAAGAGAGAC 1314
Qy 1126 AGTCTGTGAAGTTTCAGTCTCGCTCAGTGAAGCTTTGCTTCTCCCATCTTAAATG 1185
Db 1315 GTTGGGTTACCTTCTGCTCTGAGAAATGCCCAACCTTTAACGTGCGATGGCGGAGAGC 1374
Qy 1186 GCCCAACTCTTCAAGAGAGGATGAAGTCTTATAAACTGCACTGATCACCCTTTATCTCA 1245
Db 1375 GCACCTTTAACCGAGACCTCATCACCAGGTTAAGATCAAGGTCTTTTCACTCGCCGCGC 1434
Qy 1246 TTGTGTTGTAGTTCTCGTGGCCCATCAATTTGCATAGTGGCAGCTCAGCTCTCGTGAATGG 1305
Db 1435 ATGNACCCAGACAGAGTCCCTCATCTGATCGTGAAGCTTGGCTTTTGAACCC 1494
Qy 1306 AAACGAGAAATTTGCAAGGTTGGCTCAGTTAATGAGATATATCTCAAGTCCGGAAGGCA 1365
Db 1495 CTCCTGGGTCAAGCCCTTTGTACACCTTAAGCCCTCGCCCTCTTCTCCATCCGCGCC 1554
Qy 1366 AAGMAATGGCAGTGA-----AGATG 1386
Db 1555 CGTCTCTCCCTTTGAACCTCTCGTTCGACCCCGCTCGATCTCTCCCTTTATCCAGCCC 1614
Qy 1387 AAATGAGATTTTCGAGAGCTGTGATGGAAC-----GCATGAGCAACATGGA 1432
Db 1615 TCACCTCTCTTAGGCGCCGGAATTTGATCCGGACCATGAGCTTCAATACCTGATGGA 1674
Qy 1433 AAGCAGATCCAGTATCTTTTCAGATAATGAAGCCATCTCTAGATGCTAAGAAATTTCCA 1492
Db 1675 CTGGAACAGCTGTAGCCCTGAACAGCAGCGTGGCTGTGACGGCTCGCGGATTTTCGCG 1734
Qy 1493 AAATTTGAGATAAC-----AACTGATCAAGATTTAATGATGTTCTTTTCCAGCT 1543
Db 1735 CTCTGACAGTATTACCGGACGGTCAGCGATATTCTGGATTAATGTAAACCGCGGCTGA 1794
Qy 1544 AAATTTCTTACTTTCTCCATCCAGGACATGAGAAATATCATAGG----- 1589
Db 1795 CGATGCCCTCGTGTGATACAGGCTAAATTTGATATAAACAAGATGACAGCGCTACCGGT 1854
Qy 1590 -----GATATCTCAAAGTCAATAGTAGGCTGAAACCCAGCTACTTTGATTTGCA 1639
Db 1855 CACCCTTGAAGATCCGCGCGCGCGCTGAGCGCAAGTAAACAGGCGAT 1914
Qy 1640 GTTCAGTATTGAACACTGATGAGCAGAGTCA---AGAGAATGCAATTTAAACACAGGA 1696
Db 1915 GACCGCTGCCGTCAAAAATAATTGAACCGTTCATTTCCGCGAGACGCTACCGCTGTAGA 1974
Qy 1697 GGAGATCGTAAATTTAGAGGAGCGTATATACATGCATCAGCAGAAATTAAGTCTCTAGA 1756
Db 1975 TGTGAAACCCAGCCAGCGCGTGTGCGCAGAGGTTACCGGTCCCGCTCTGTCTGCG 2034
Qy 1757 T---GAAAAACAAGTATATTTGGACAGAGAAATAAAGGGAAATGAACCTGTTGAATAA 1813
Db 2035 TCTGTATATTTCCGCGGCTCGGCTCCGCTCTTCTCAACGGTGTGATGCTTGGGACGCC 2094
Qy 1814 TATCACTAATGATCTGAGCGTGAAGGATTTGGGAACATTTCTCAGACATTTGAATAATCAC 1873

Db 2095 GGCGCGCATTTGCGGGATGCCAGAAAGGTGGTTCTGTCTCGCGCGGCCCATCGCTGATGA 2154
Qy 1874 TTTACTCAAGGTGCCAGAAAGTGTCTCGTACATGGGAAATGGACCAACG----- 1923
Db 2155 AATCCTCTATGCGCGCAACTGTGTGGCGTGCAGGAAATCTTTAACGTGCGCGCGCGCA 2214
Qy 1924 -----ATCTGGGCTCCAACATGACCATCGGGCTGTGAAACAGCAGAGGTGAATTCACAGG 1978
Db 2215 GCGGATTTGCGGCTCTGCGCTTTCGCGAGCGAGTCCGTATACCGAAAGTGGATNAATTTTGG 2274
Qy 1979 CACCTACATCACAGCCGTAAACAGCACATCAAAATGAGATCA---AAGAGTCAACCACTGCAT 2036
Db 2275 CCGCGGCAACGCCCTTTGTAAACGAAGCAACAGCTGAGTCAAGCCAGCTCTCGACGGCGC 2334
Qy 2037 GGGACACAAAACACCATC-----AAACAAGAGGACCCAGCCCACTTTTGGCT 2082
Db 2335 GGCTATCGATATGCCAGCGGCGCTCTGAAGTACTGTGATCGCAGACAGCGCGCAAC 2394
Qy 2083 TCACCGTCAATTTGAAAGTTTTCAGAGTCCACCACTGTCTTCAAGGCGC----- 2130
Db 2395 ACCGGATTTGCTGCTTCTGACCTGCTCTCCAGGCTGAGCAGCGCCCGGATTCACAGGT 2454
Qy 2131 -----AGTCTTCATAGACAGCAATGGAAGGAGTCTCTGAAGACCAT 2173
Db 2455 GATCCTGTGACGCTGATGCTGACATTTGCCCGCAAGTGGCGGAGCGGTAGAACGTCA 2514
Qy 2174 GTGCTGCTGCGGTCAAGTGTTAATGATCTGATGCTGATGCTGAAAGCTA---CCAGGGTCT 2231
Db 2515 ACTGGCGGAATCTCGCGCGCGGACACCGCCCGCAGGCGCTGAGCGCAGTCTGCTGAT 2574
Qy 2232 GGCATCAACATCTTCACTGCTGCTGCGCACACAGAAAGAGTGTGATGAGTGAACAAGTCT 2291
Db 2575 TGTGACCAAAAGATTTAGCGCAGTCCGTGCGCATCTCTAATCAGTATGCGCGGAACACTT 2634
Qy 2292 CCTGGACTCCAGGTGAAGAGGATAGAGGCCCTCTTGACAA----- 2335
Db 2635 AATCATCAGACGCGCAATGCGCGGATTTGGTGGATGCGGATTAACAGCGAGGCTCGGT 2694
Qy 2336 -----AATGGTATACAGGCTTTCCAGGCTCTAATAGTACTTCAGGCTCTTAAAGG 2385
Db 2695 ATTTCTCGCGAGCTGTGCGCGGAATCCGCGGTGATTAACGCTTCGGAACCAACCATGT 2754
Qy 2386 TGATCGGGGGATCTCTGTTTACTGAGTTTCAGAGTTCCAGGATTCAGGACCAATGGGGAAGA 2445
Db 2755 TTAAACGACCTATGGCTATCTGCTACTGTTTCCAGCCTGGGTTAGCGGATTTCCAGAA 2814
Qy 2446 CCGGAGACGACGACTTAATGACAAAGG-----CCAGAAAGGA 2486
Db 2815 ACGGATGACGCTTCAGGAACCTGTGAAAGCGGGGCTTTTCCGCTCTGGCATCAACCATTTGA 2874
Qy 2487 GAAAAAGGAGTGGAAAGCATGCAAAAGCAATCTAATACAGTCCGACTGTTGGTGGCAGC 2546
Db 2875 AACATTGGCGGCGAGAACGCTGTGACCGGCCATTAATAATCGCTGACCTTGGCGTAAA 2934
Qy 2547 GGCCCTCACGAAGCGCAGATG-----GAGATTTTTCAGAAAGCGCAGTGGGTACCGTG 2600
Db 2935 CGCCCTCAAGAGGCAAGCATGAGCACTGAAACACTCTCAGCGCTCGCGGATCAATTCG 2994
Qy 2601 TGTGACGACCGCTGGGAACCTGCGTGGAGGACTG---GTGCTGTGAGGAGCTTGGGATCAA 2659
Db 2995 TTAACTCGAGGATCGGCCAGCCACGAGTGGTATAGTATAGNATAGATTTTATTTAGA 3054
Qy 2660 AGGTGTTCAAAGTGTGATAGCGAGCTTATTTTGGAAAGGTACGGGTCCAATATGGCT 2719
Db 3055 GCATAGGAGGGAGTCAAGGGCAGTAGAGGCAAGAAAGAGAGAGAGGGGGAGAG 3114
Qy 2720 GAATGAAGTATTTTGTTCGGGAAAGATC---ATCCATTTGAAGAGTGCAGAAATTAGACAG 2777
Db 3115 AAGAGAGTAGAGGTGTAGAGGCTGGCTTGGAGCAGCTGGAGAGAGCAGGGAAGAA 3174
Qy 2778 TGGGTTGTGAGAGCTGTTTCGCAC-----GACGAAGATCTCGGGGTCACT 2823
Db 3175 TGGGAGAGAGGGNAGAAATTTGGAACAGAGAAAGAGAGAGAGATCAAGAGAGCAAGG 3234

Db 5395 AGCTTAAGTAACGCCATTTTGCAGGCAATGGAATAATACATACATGGAATAGAGAGTT 5454
Qy 4674 CAGATCAAGGTCAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGGT 4733
Db 5455 CAGATCAAGGTCAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGGT 5514
Qy 4734 AAGCAGTTCTGCCCCGGCTCAGGCCAAGAACAGATGGAACAGCTGAATATGGGCCAAA 4793
Db 5515 AAGCAGTTCTGCCCCGGCTCAGGCCAAGAACAGATGGAACAGCTGAATATGGGCCAAA 5574
Qy 4794 CAGGATATCTGTGGTAAGCAGTTCTGCCCCGGCTCAGGCCAAGAACAGATGTCGCCCA 4853
Db 5575 CAGGATATCTGTGGTAAGCAGTTCTGCCCCGGCTCAGGCCAAGAACAGATGTCGCCCA 5634
Qy 4854 GATCGGTCAGGCCCTCAGCAGCTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCAA 4913
Db 5635 GATCGGTCAGGCCCTCAGCAGCTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCAA 5694
Qy 4914 GGACCTGAATGACCTGTGCTTATTTGAATCAATCAATCAATCAATCAATCAATCAATCAAT 4973
Db 5695 GGACCTGAATGACCTGTGCTTATTTGAATCAATCAATCAATCAATCAATCAATCAATCAAT 5754
Qy 4974 TTGCGGCTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTTCGGGGCGC 5033
Db 5755 TTGCGGCTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTTCGGGGCGC 5814
Qy 5034 CAGTCTCCGATTCAGTCCGCGGCTACCGGTGATCCATTAATCAATCAATCAATCAATCAAT 5093
Db 5815 CAGTCTCCGATTCAGTCCGCGGCTACCGGTGATCCATTAATCAATCAATCAATCAATCAAT 5874
Qy 5094 TGCATCCAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5153
Db 5875 TGCATCCAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5934
Qy 5154 CGTACGCGGGCTCTTTCAATTGG 5177
Db 5935 CGTACGCGGGCTCTTTCAATTGG 5958

RESULT 3
US-09-380-190A-29
; Sequence 29, Application US/09380190A
; Patent No. 6410220
; GENERAL INFORMATION:
; APPLICANT: NATURE TECHNOLOGY CORPORATION, ET AL.
; TITLE OF INVENTION: SELF-ASSEMBLING GENES, VECTORS AND USES
; THEREOF
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MUETING, RAASCH & GEBHARDT, P.A.
; STREET: 119 NORTH FOURTH STREET, SUITE 203
; CITY: MINNEAPOLIS
; STATE: MINNESOTA
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/380,190A
; FILING DATE: 26-Aug-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/03918
; FILING DATE: 28-FEB-98
; ATTORNEY/AGENT INFORMATION:
; NAME: MUETING, ANN M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 228. 00010201
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 612-305-1217
TELEFAX: 612-305-1228
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5594 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-09-380-190A-29

Query Match 31.5%; Score 1629.4; DB 3; Length 5594;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1654; Conservative 0; Mismatches 21; Indels 2; Gaps 1;

Qy 3371 AAGCTTCAAGCTCGCCGCAAGCAGCTCAGGGCCAAAGGGCTGCTAAAGGAAGCGGAACAGT 3430
Db 3342 AAGCTTCAAGCTCGCCGCAAGCAGCTCAGGGCCAAAGGGCTGCTAAAGGAAGCGGAACAGT 3401
Qy 3431 AGAAAGCCAGTCCGCAAGAACCGGTGCTGACCCCGGATGAATGTGAGCTTACCTGGGTATCT 3490
Db 3402 AGAAAGCCAGTCCGCAAGAACCGGTGCTGACCCCGGATGAATGTGAGCTTACCTGGGTATCT 3461
Qy 3491 GGACAAGGGAAAACGCAAGCGCAAGAAAGCAAGGTAGCTTGCAGTGGGGCTTACATGCG 3550
Db 3462 GGACAAGGGAAAACGCAAGCGCAAGAAAGCAAGGTAGCTTGCAGTGGGGCTTACATGCG 3521
Qy 3551 GATAGCTAGACTGCGGCGGTTTTATGGACAGCAAGCAACCGGAATTCACCTGGGGCGC 3610
Db 3522 GATAGCTAGACTGCGGCGGTTTTATGGACAGCAAGCAACCGGAATTCACCTGGGGCGC 3581
Qy 3611 CCTCTGTTAAGTTGGGAAGCCCTGCAAGTAATCTGATGGCTTTCTTGGCGCCCAAGGA 3670
Db 3582 CCTCTGTTAAGTTGGGAAGCCCTGCAAGTAATCTGATGGCTTTCTTGGCGCCCAAGGA 3641
Qy 3671 TCTGATGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCGTTTGCATGA 3730
Db 3642 TCTGATGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCGTTTGCATGA 3701
Qy 3731 TTGAACAAGATGGATTGCACGACAGTTCTCCGGCGGCTTGGGTGGAGAGGCTATTTCGGCT 3790
Db 3702 TTGAACAAGATGGATTGCACGACAGTTCTCCGGCGGCTTGGGTGGAGAGGCTATTTCGGCT 3761
Qy 3791 ATGACTGGGCACACAGACAAATCCGGCTGCTCTGATGCGCGCTGTTCGGGCTGTCCAGCGC 3850
Db 3762 ATGACTGGGCACACAGACAAATCCGGCTGCTCTGATGCGCGCTGTTCGGGCTGTCCAGCGC 3821
Qy 3851 AGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTCCCTGAACTGACAGG 3910
Db 3822 AGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTCCCTGAACTGACAGG 3881
Qy 3911 ACAGGGCAGCGCGCTATCGTGGCTGGCCAGCAGCGGGGTTCTTGGCAGCTGTGCTCG 3970
Db 3882 ACAGGGCAGCGCGCTATCGTGGCTGGCCAGCAGCGGGGTTCTTGGCAGCTGTGCTCG 3941
Qy 3971 ACCTTGTCACTGAAGCGGAGGACCTGGCTGCTATTGGGGCGAAAGTCCCGGGCAGGATC 4030
Db 3942 ACCTTGTCACTGAAGCGGAGGACCTGGCTGCTATTGGGGCGAAAGTCCCGGGCAGGATC 4001
Qy 4031 TCCTGTCACTCACTTCTCTCCCGGAGAAAGTATCCATCATGCTGATGCAATGCGCGC 4090
Db 4002 TCCTGTCACTCACTTCTCTCCCGGAGAAAGTATCCATCATGCTGATGCAATGCGCGC 4061
Qy 4091 GGCTGCATACGCTTGTATCCGGCTACCTGCCCATTCGACCACCAAGCGGAACATCGCATCG 4150
Db 4062 GGCTGCATACGCTTGTATCCGGCTACCTGCCCATTCGACCACCAAGCGGAACATCGCATCG 4121
Qy 4151 AGCGACGACGCTACTCGGATGGAAGCCGCTTTGTGCGATCAGGATGATCTGACCAAGAGC 4210
Db 4122 AGCGACGACGCTACTCGGATGGAAGCCGCTTTGTGCGATCAGGATGATCTGACCAAGAGC 4181
Qy 4211 ATCAGGGGCTCGCGCCAGCCGAACTGTTGCGCAGGCTCAAGGCGCGCATGCCCGAGCGC 4270

```
Db 4182 ATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGCGCGCATGCCCGAGCGG 4241
Qy 4271 AGGATCTCGTGTGACCCATCGCATGCTGCTTCCGCAATATCATGTGGAATAATGGCC 4330
Db 4242 AGGATCTCGTGTGACCCATCGCATGCTGCTTCCGCAATATCATGTGGAATAATGGCC 4301
Qy 4331 GCTTTCTCGAATTCATCCAGCTGTGCGCGCTGGGTGTGCGGACCGCTATCAGGACATAG 4390
Db 4302 GCTTTCTCGAATTCATCCAGCTGTGCGCGCTGGGTGTGCGGACCGCTATCAGGACATAG 4361
Qy 4391 CGTTGGCTACCCGTGATATGCTGAAGAGCTTGGCGGCAATGGGCTGACCGCTTCTCTCG 4450
Db 4362 CGTTGGCTACCCGTGATATGCTGAAGAGCTTGGCGGCAATGGGCTGACCGCTTCTCTCG 4421
Qy 4451 TGCTTTACGGTATGCGCGCTCCGATTCGAGCGGATCGCTTCTATCGCCTTCTTGACG 4510
Db 4422 TGCTTTACGGTATGCGCGCTCCGATTCGAGCGCATCGCTTCTATCGCCTTCTTGACG 4481
Qy 4511 AGTTCTTCTGAGCGGACTCTGGGTTTCGATAAATAAAGATTTTA--TTTAGTCTCCA 4568
Db 4482 AGTTCTTCTGAGCGGACTCTGGGTTTCGATAAATAAAGATTTCGACCGGACCCACCTCCA 4541
Qy 4569 GAAAGAGGGGGAATGAAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAAGTAACGCC 4628
Db 4542 GAAAGAGGGGGAATGAAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAAGTAACGCC 4601
Qy 4629 ATTTTGAAGGATGGAATAATACATACTGAGAAATAGAGAAGTTTCAGATCAAGGTGAG 4688
Db 4602 ATTTTGAAGGATGGAATAATACATACTGAGAAATAGAGAAGTTTCAGATCAAGGTGAG 4661
Qy 4689 AACAGATGGAACAGCTGAAATAGGCGCAACAGGATATCTGTGTAGCAGTTCCTGCC 4748
Db 4662 AACAGATGGAACAGCTGAAATAGGCGCAACAGGATATCTGTGTAGCAGTTCCTGCC 4721
Qy 4749 CGGCTCAGGCGCAAGAACAGATGGAACAGCTGAAATATGGGCAAAACAGGATATCTGTGT 4808
Db 4722 CGGCTCAGGCGCAAGAACAGATGGAACAGCTGAAATATGGGCAAAACAGGATATCTGTGT 4781
Qy 4809 AAGAGTTTCTGCGCGCGCTCAGGCGCAAGAACAGATGTTCCCGAGATGCGGTCCAGGCC 4868
Db 4782 AAGAGTTTCTGCGCGCGCTCAGGCGCAAGAACAGATGTTCCCGAGATGCGGTCCAGGCC 4841
Qy 4869 TCAGAGTTTCTAGAGAACCATCAGATGTTTCCAGGTTCCAGGACCTGAAATGACC 4928
Db 4842 TCAGAGTTTCTAGAGAACCATCAGATGTTTCCAGGTTCCAGGACCTGAAATGACC 4901
Qy 4929 CTGTGCTTATTTGAACCTAACCAATCAGTTGCTGCTTCTGCTTCTGCTGCGCGCTTCTGC 4988
Db 4902 CTGTGCTTATTTGAACCTAACCAATCAGTTGCTGCTTCTGCTTCTGCTGCGCGCTTCTGC 4961
Qy 4989 TCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGCGGCGCATGCTCCCGAT 5045
Db 4962 TCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGCGGCGCATGCTCCCGAT 5018
```

RESULT 4

```
US-09-897-511A-9
; Sequence 9, Application US/09897511A
; Patent No. 6852510
; GENERAL INFORMATION:
; APPLICANT: Brenel, Robert
; APPLICANT: Miller, Linda
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors
; FILE REFERENCE: GALA-06416
; CURRENT APPLICATION NUMBER: US/09/897,511A
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,925
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
```

```
; LENGTH: 5130
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-897-511A-9

Query Match      30.7%; Score 1588.6; DB 3; Length 5130;
Best Local Similarity 66.1%; Pred. No. 0;
Matches 3007; Conservative 0; Mismatches 754; Indels 786; Gaps 18;

Qy 1 TTTGAAAGACCCCAACCGTAGGTGGCAAGCTAGCTTAAAGTAAACGCCACTTTTGCAGGCAAT 60
Db 1 TTTGAAAGACCCCAACCGTAGGTGGCAAGCTAGCTTAAAGTAAACGCCACTTTTGCAGGCAAT 60
Qy 61 GGAATAATACATTAATCAGATAAGAAAGTTTCAGATCAAGTCAAGGACCAAGAAACAGC 120
Db 61 GGAATAATACATTAATCAGATAAGAAAGTTTCAGATCAAGTCAAGGACCAAGAAACAGC 120
Qy 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTTCTGCGCCCGCTCAGGGCCCAAGAAC 180
Db 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTTCTGCGCCCGCTCAGGGCCCAAGAAC 180
Qy 181 GATGAGACAGCTGAGTGCATGCGGCCAAACAGGATATCTGTGTAAAGCAGTTCCTGCCCGG 240
Db 181 GATGAGACAGCTGAGTGCATGCGGCCAAACAGGATATCTGTGTAAAGCAGTTCCTGCCCGG 240
Qy 241 CTGCGGGCCAAAGAACAGATGTTCCCGAGATGCGGTCCAGCCCTCAGCAGTTCCTAGTGAA 300
Db 241 CTGCGGGCCAAAGAACAGATGTTCCCGAGATGCGGTCCAGCCCTCAGCAGTTCCTAGTGAA 300
Qy 301 TCATCAGATGTTTCCAGGGTGCCCAAGGACCTGAAATGACCCCTGTACCTTATTTGAAAC 360
Db 301 TCATCAGATGTTTCCAGGGTGCCCAAGGACCTGAAATGACCCCTGTACCTTATTTGAAAC 360
Qy 361 TAACCAATCAGTTTCGCTTCTGCTTCTGTTGCGCGCTTCCGCTCTCCGAGCTCAATAAA 420
Db 361 TAACCAATCAGTTTCGCTTCTGCTTCTGTTGCGCGCTTCCGCTCTCCGAGCTCAATAAA 420
Qy 421 AGAGCCCAACCCCTCACTCGCGCGCCAGTCTTCCGATAGACTGCGTCCCGGGGTAC 480
Db 421 AGAGCCCAACCCCTCACTCGCGCGCCAGTCTTCCGATAGACTGCGTCCCGGGGTAC 480
Qy 481 CCGTATTTCCCAATAAAGCCTTCTGCTGTTTGCATCCGAAATCGTGCTCTCGCTGTTCTTG 540
Db 481 CCGTATTTCCCAATAAAGCCTTCTGCTGTTTGCATCCGAAATCGTGCTCTCGCTGTTCTTG 540
Qy 541 GGAGGGTCTCTCTGAGTGAATGACTACCCAGACCGGGGGTCTTTTCAATTTGGGGGCTCGT 600
Db 541 GGAGGGTCTCTCTGAGTGAATGACTACCCAGACCGGGGGTCTTTTCAATTTGGGGGCTCGT 600
Qy 601 CCGGGATTTGGAGACCCCTGCGCGGAGACCAACCCAGCCACCGGGAGGTAACTGGCC 660
Db 601 CCGGGATTTGGAGACCCCTGCGCGGAGACCAACCCAGCCACCGGGAGGTAACTGGCC 660
Qy 661 AGCAACTTATCTGTGCTGCTGCGGATTCCTAGTGTCTATGTTTGTATGCTGCTGCTGCTG 720
Db 661 AGCAACTTATCTGTGCTGCTGCGGATTCCTAGTGTCTATGTTTGTATGCTGCTGCTGCTG 720
Qy 721 TCTGTACTAGTTAGTAACTAGCTCTGTATCTGCGGACCCCGTGGTGAACCTGACGAGTT 780
Db 721 TCTGTACTAGTTAGTAACTAGCTCTGTATCTGCGGACCCCGTGGTGAACCTGACGAGTT 780
Qy 781 CTGAACACCCCGCGCAACCCCTGCGGAGACGCTCCAGGAGATTTTGGGGCCCTTTTGTGG 840
Db 781 CTGAACACCCCGCGCAACCCCTGCGGAGACGCTCCAGGAGATTTTGGGGCCCTTTTGTGG 840
Qy 841 CCGGACCTGAGGAGGAGTGCATGCGAATCCAGCCCGCTCAGGATATGCTGCTCTGCT 900
Db 841 CCGGACCTGAGGAGGAGTGCATGCGAATCCAGCCCGCTCAGGATATGCTGCTCTGCT 900
Qy 901 AGGAGACGAGAACCTAAACAGATTTCCCGCCCTCCGCTCTGAAATTTTTCGTTTGGAA 960
Db 901 AGGAGACGAGAACCTAAACAGATTTCCCGCCCTCCGCTCTGAAATTTTTCGTTTGGAA 960
```

Db 901 AGGAGACGAGAACCTAAACACAGTTCCCGCCTCCGCTGTAATTTTGTCTTTTCGGTTTGGAA 960
Qy 961 CCGAAGCCGCGCTCTTCTGCTGACGACCAAGCTTGGCTGCGAGTTCGACTCTAGAGGA 1020
Db 961 CCGAAGCCGCGCTCTTCTGCTGACGACCAAGCTTGGCTGCGAGTTCGAGTTCGAGGA 1020
Qy 1021 TCMAATTCGGCAGGAGTAATCGGTGCTGCGCTCTTTAGGACATATGAAGTATGCAACGT 1080
Db 1021 CC-----CTGCACCCGAAACATGGAGAACCAACATCAGGATTCCT 1059
Qy 1081 GGGATGACTTTCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTGCG 1140
Db 1060 AGGACCCCTGCTGTTTACAGCGGGGTTTCTTGTGTGACAGAACTCTCACATACC 1119
Qy 1141 ATGCTGCTCAGTGACAGCTTTGCTTCTCCCAATCCCAATTAATAATGGGCCAACTCTTCAAG 1200
Db 1120 ACAGAGTCTAGAGTCTGCTGAGTCTCTCAATTTTCTAGGGGAGCACCCACGTGTCC 1179
Qy 1201 AGAGGATGAAGTCTTATAAAATGCACTGATGATCACCCTTATCTCATTTGTGTTGTAGTTC 1260
Db 1180 TGGCCAAATTCGAGTCCCAACCTCCAATCACTCACCAACCTCTTGTCTCCTCAATTTG 1239
Qy 1261 TCGTGCCCATCATTTGGCATAGTGGCAGCTCAGCTCCTGAAATGGGAAACGAAGATTGCA 1320
Db 1240 TCCTGGCTATC-----GCTGGATGTCTCGCGCTTTTATC 1276
Qy 1321 CGGTTGGCTCAGTTAATGACAGATATCTCCAAGTCCGGAAGGCAAGAAATGGCAGTG 1380
Db 1277 ATATTCTCTTCATCTCTGCTATGCTCATCTTCTTGTGTTGTTCTTCTGCACTACCA 1336
Qy 1381 AAGATGAATGAGATTTTCGAGAGCTGTGATGGAACGATGAGCAACATGAGAAAGCAGAA 1440
Db 1337 GGTATGTTGGCCGTTTGTCTCTACTTCCAGGAA-----CATCAATACCAACGACGAGCA 1393
Qy 1441 TCCAGTATCTTTCAGATTAATGAAGCCATCTCTAGATGCTTAAGATTTCCAAATTTCA 1500
Db 1394 TGCAGAGCTCGCAGATTCCTGCTCAAGAACTCTATGTTTCCCTCTT-----GTTGCT 1448
Qy 1501 GCATAACAACATGATCAAGATTTAATGATGTTCTTTTCCAGCTAAATTCCTTACTTCTCT 1560
Db 1449 GTACAAAACCTTCGAGCGGAACTGCACCTTGTATTTCCCATCATCTCTGGGCTTTCG 1508
Qy 1561 CCATCCAGGAACATGAGAAATCATAGGGGATATCTCCAAGTCAATAGTAGGTCTGAACA 1620
Db 1509 CAAGATTCCTATGGAGTGGGCTCAGTCCCTTCTCTCTCG----- 1549
Qy 1621 CCACAGTACTTGAATTTGAGTTTCAATGATTTGAACACTGAATGGCAGAGTCCAGAGATG 1680
Db 1550 -CTCAGTTTACTAGTGCCATTTGTTTCACTGTTGTC---GTAGGGCTTTTCCCTCCTGTTG 1605
Qy 1681 CATTTAAACAACAGAGGAGATGCTTAATTTAGAGGAGCGTATATACATGCATCAGCAG 1740
Db 1606 GCTTTCAGTTATAG-GATGATGTTGTTATTTGGGGCCCAAGTCTGTACAACTCTTGATC 1664
Qy 1741 AAATTAAGTCTCTAGATGAAACAAAGTATATTTTGGAAACAGAAATAAAGGGGAAATGA 1800
Db 1665 CCTTTTACCTCTATTAACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCCCTAATA 1724
Qy 1801 AACTGTTGAATATCACTAATGATCTGAGGCTGAAGGATTTGGAAACATTTCTCAGACAT 1860
Db 1725 AAC-----CAACGT 1734
Qy 1861 TGAATAATATCATTTTACTCCAAGTGCAGAAAGTGTCTGCTGACTGGGAAATGGACCA 1920
Db 1735 TGGGCTACTCCCTTAACCTTCAATGATGATGATGATGATGATGATGATGATGATGATG 1781
Qy 1921 ACATGCTGGGCTCCAACATGATCCATCCGCGGCTGTGAACAGCAGAGAGTGAATTCAGGCA 1980
Db 1782 ----- 1781
Qy 1981 CCTACATCAGCGGTAAACAGCCATCAATAGATGATCAAGAGTCAACCTGATGAGGA 2040
Db 1782 -----TACTTTACCGCAAGAACATATTGTACTATAAATAATCAAGCAATGTTTTCG- 1829

Qy 2041 CACAAAACACATCAACAAAGAGGACCCAGGCCACCTTTTGGCTTCCCGTCAATTTGGAAGT 2100
Db 1830 -----AAAACCTGCTCTTAATAATAGACCTTATTGATTTGGAAGATATGTCAGAGACTTTGTTGGT 1884
Qy 2101 TTTTCAGAGTCCACCACTGTCTTTCAGCGGCCAGTGTCTTCATAGACAGGAATGGGAAGGAG 2160
Db 1885 CTTTTTGGGCTTTGTGCTCCCTTTTACAAATGTGGCTATCTGCTCTTAATGCCCTTTATAT 1944
Qy 2161 TCCTGAAGACCATGTGCTGCTGCGGTCAAGTGTAAATGACATTTGGTGATGACTGGAAG 2220
Db 1945 GCATGTATACAATCT----- 1959
Qy 2221 CTACAGGCTCGGATCAACATCTTCACTGCTGCGCACACAGAGAGTGAAGTGAAGT 2280
Db 1960 -----AAGCAGGCTTTTCACTTTCTCGCAACTTACAGGCTTTCTGTGTA 2005
Qy 2281 ACCAAGTCTCTCTGAGTCCAGGTGAAGAGGAGATAGAGGCCCTCTGCGCAAAATGG 2340
Db 2006 AACAAATATCGAACCTTTTACCCCGTTGCCCGCA----- 2039
Qy 2341 TATACCAGGCTTTCCAGGTCTAATAGGTACTCCAGGTCTTAAAGGTGATCGGGGATCT 2400
Db 2040 ----- 2039
Qy 2401 CTGTTTACCTGGAGTTGAGGATTTCCAGGACCAATGGGGAAGACCGGGAAGCCAGAC 2460
Db 2040 ----- 2039
Qy 2461 TTAATGACAAAAAGGCCAGAGGAGAAAAAGGGAGTGGAGCATGCAAGACAAATCTA 2520
Db 2040 -----ACGTCAG 2047
Qy 2521 ATACAGTCCGACTGCTGGTGGGAGCGCCCTCACGAAGGAGGAGTGGAGATTTTTCACG 2580
Db 2048 GTCTCTCCAAAGTGTTCGACGCAACCCCACTGATGGGGCTTGGCTATCGGCCATA 2107
Qy 2581 AAGGCCAGTGGGTPACGGTGTGACGACCGCTGGGAACTGCGTGGGAGTGGTCTGCT 2640
Db 2108 GCGCATGCGCGGACCTTTTGTGGCTCTCTGCG----- 2141
Qy 2641 GCAGGACTTGGGATCAAAAGGTGTTCAAAGTGTGCTAAGCGAGCTTATTTTGGAAAAG 2700
Db 2142 -----ATCCATATGCGGAACCTCTAGCAGCTTGT----- 2171
Qy 2701 GTACGGGTCCAAATATGCTGAATGAAGTATTTTGTTCGGGAAAGAGTCAATCCATGAAG 2760
Db 2172 ----- 2171
Qy 2761 AGTCAGAAATTAGACAGTGGGTTGTAGAGCTGTCGACGACGAAAGATGCTGGGGTC 2820
Db 2172 -----TTTGTCTGCGAGGGGTCTGGAGCGAA 2197
Qy 2821 ACTTTGACCTACATAATGATCATATTTTTCATTCACATTTTTTAAACTGTTATAAGTG 2880
Db 2198 ACTTATCGGC----- 2207
Qy 2881 ATTTTTTCTTTGCTTCACTAAAATCAGCTTAATTAATATTTAAGAAACTAAGAAATTTT 2940
Db 2208 ----- 2207
Qy 2941 ATCCACAGAAAAAGGAATATTTAAAAATCACTGGATAAACATATAAAATAGCTTTCATATTT 3000
Db 2208 -----ACCGCAACTCT 2219
Qy 3001 GCTTCAAAATACAGAACCAATTTCACTTCTCTAGTGTATTAAGTGGCTCGTGGCGAAATG 3060
Db 2220 GTTGTCTCTCTCGAAAAATACACCTCTCTTCCATGGCTGTAGGTGTGCTGCGCAACTGG 2279
Qy 3061 ATCCCTCAGATATAGTACTTTCGCTTTGCTATGAGGAGGGGAAATGATGCTTATGCG 3120
Db 2280 ATCCCTCAGATATAGTACTTTCGCTTTGCTATGAGGAGGGGAAATGATGCTTATGCG 2339

QY 3121 AATACTCTTGTAGTCTTGCAACATGGTAACGATGAGTTAGCAACATGCTTTACAGGAGA 3180
DB |||||
QY 2340 AATACACTTGTAGTCTTGCAACATGGTAACGATGAGTTAGCAACATGCTTTACAGGAGA 2399
DB |||||
QY 3181 GAAAGACACCGTGTGATCCGATTCGATGGAAGTAAAGTGGTACGATCGTCTTATTAGG 3240
DB |||||
QY 2400 GAAAGACACCGTGTGATCCGATTCGATGGAAGTAAAGTGGTACGATCGTCTTATTAGG 2459
DB |||||
QY 3241 AAGGCAACAGACGGTCTGACATGATTTGGAGCAACCACTGAATTCGCGATTCGAGAT 3300
DB |||||
QY 2460 AAGGCAACAGACGGTCTGACATGATTTGGAGCAACCACTGAATTCGCGATTCGAGAT 2519
DB |||||
QY 3301 -ATTGTAATTTAAGTCCCTAGCTCGATACAGCAAAACGCCA--TTTGACCAATTCACCACTT 3357
DB |||||
QY 2520 AATTGTAATTTAAGTCCCTAGCTCGATACAGCAAAACGCCAATTTTGACCAATTCACCACTT 2579
DB |||||
QY 3358 GGTGTGCACT - CCAAGCTTTCAGCTGCCGCAAGCACTCAGGGCGCAAGGGCTGCTAAA 3415
DB |||||
QY 2580 GGTGTGCACTTCCAAAGCTTCAGCTGCCGCAAGCACTCAGGGCGCAAGGGCTGCTAAA 2639
DB |||||
QY 3416 GGAAGCGGAACACGTAAGAACCGAGTCCGAGAAACGCTGCTGAGCCCGGATGAATGTCA 3475
DB |||||
QY 2640 GGAAGCGGAACACGTAAGAACCGAGTCCGAGAAACGCTGCTGAGCCCGGATGAATGTCA 2699
DB |||||
QY 3476 GCTACTGGGCTATCTGGACAAGGGGAAAACGCAAGCGCAAGAGAAAGCAGGTAGCTTGCA 3535
DB |||||
QY 2700 GCTACTGGGCTATCTGGACAAGGGGAAAACGCAAGCGCAAGAGAAAGCAGGTAGCTTGCA 2759
DB |||||
QY 3536 GTGGCTTTACATGGCGATAGCTAGCTGGGCGGTTTTATGACAGCAAGCAAGCGGAAT 3595
DB |||||
QY 2760 GTGGCTTTACATGGCGATAGCTAGCTGGGCGGTTTTATGACAGCAAGCAAGCGGAAT 2819
DB |||||
QY 3596 TGCCAGCTGGGGCGCCCTCTGGTAAGTTGGAGCCCTGCAAGTAACATGGATGGCTT 3655
DB |||||
QY 2820 TGCCAGCTGGGGCGCCCTCTGGTAAGTTGGAGCCCTGCAAGTAACATGGATGGCTT 2879
DB |||||
QY 3656 TCTTGCCGCCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGATGAG 3715
DB |||||
QY 2880 TCTTGCCGCCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGATGAG 2939
DB |||||
QY 3716 GATCGTTTTCGATGATGAACAAGATTTGACGCAAGTTTCTCGGCCGCTTGGGTGG 3775
DB |||||
QY 2940 GATCGTTTTCGATGATGAACAAGATTTGACGCAAGTTTCTCGGCCGCTTGGGTGG 2999
DB |||||
QY 3776 AGAGCTATTTCGCTATGACTGGGCAACAAGACATCGCTGCTGATCGCGCGCTGT 3835
DB |||||
QY 3000 AGAGCTATTTCGCTATGACTGGGCAACAAGACATCGCTGCTGATCGCGCGCTGT 3059
DB |||||
QY 3836 TCCGCTCTGACGCGAGGGCGCCCGGTTCTTTTGTCAAGACGACCTGTCCGCTGCC 3895
DB |||||
QY 3060 TCCGCTCTGACGCGAGGGCGCCCGGTTCTTTTGTCAAGACGACCTGTCCGCTGCC 3119
DB |||||
QY 3896 TGAATGAATGACGAGCAGGCGCGGCTATCGTGGCTGGCCACGAGCGGCGTTCCTT 3955
DB |||||
QY 3120 TGAATGAATGACGAGCAGGCGCGGCTATCGTGGCTGGCCACGAGCGGCGTTCCTT 3179
DB |||||
QY 3956 GCGCAGCTGTGCTGACGCTGTGCTGAAGCGGGAAGGACATGCGTGTCTATTGGCGGAAG 4015
DB |||||
QY 3180 GCGCAGCTGTGCTGACGCTGTGCTGAAGCGGGAAGGACATGCGTGTCTATTGGCGGAAG 3239
DB |||||
QY 4016 TGCGGGGCGAGATCTCTGTCATCTCACCTTGCTTCTGCGGAGAAAGTATCCATATGG 4075
DB |||||
QY 3240 TGCGGGGCGAGATCTCTGTCATCTCACCTTGCTTCTGCGGAGAAAGTATCCATATGG 3299
DB |||||
QY 4076 CTGATGCAATGCGGCGGTGATACGCTTGATTCGGCTTACCTGGCCATTTCGACCACCAAG 4135
DB |||||
QY 3300 CTGATGCAATGCGGCGGTGATACGCTTGATTCGGCTTACCTGGCCATTTCGACCACCAAG 3359
DB |||||
QY 4136 GGAACATCGATTCGAGCGAGCAGTACTCGATGGAAGCGGCTTGTGATCAGGATG 4195
DB |||||
QY 3360 GGAACATCGATTCGAGCGAGCAGTACTCGATGGAAGCGGCTTGTGATCAGGATG 3419
DB |||||
QY 4196 ATCTGGAGGAAGAGCATCAGGGGCTCGCGCCAGCGGCACTGTTCGCCAGGCTCAAGGGCG 4255

DB 3420 ATCTGGAGGAAGACATCAGGGCTCGCGCCAGCGCAACTCTTCCCGAGGCTCAAGGGCG 3479
QY 4256 GCATGCCCGAGCGGAGGATCTCGTGTAGCCCATGGGATGCCTGTTCGCGAATATCA 4315
DB 3480 GCATGCCCGAGCGGAGGATCTCGTGTAGCCCATGGGATGCCTGTTCGCGAATATCA 3539
QY 4316 TGGTGGAAATGGCCGCTTTCTGGATTTCATCGACTGTGGCCGCTGGGTGTGGCGGACC 4375
DB 3540 TGGTGGAAATGGCCGCTTTCTGGATTTCATCGACTGTGGCCGCTGGGTGTGGCGGACC 3599
QY 4376 GCTATCAGGACATAGCCGTGGCTACCCGCTGATATTGCTGAAGAGCTTGGCGCGCAATGG 4435
DB 3600 GCTATCAGGACATAGCCGTGGCTACCCGCTGATATTGCTGAAGAGCTTGGCGCGCAATGG 3659
QY 4436 CTGACCGCTTCTCGTCTCTTACGATTCGCGCTCCGATTCGAGGCGCATCGCTTCT 4495
DB 3660 CTGACCGCTTCTCGTCTCTTACGATTCGCGCTCCGATTCGAGGCGCATCGCTTCT 3719
QY 4496 ATCGCTTCTTGACGAGTCTTCTGAGCGGAGCTCTGGGGTTCGATA 4542
DB 3720 ATCGCTTCTTGACGAGTCTTCTGAGCGGAGCTCTGGGGTTCGATA 3766

RESULT 5
US-08-352-990-1
; Sequence 1, Application US/08352990
; Patent No. 5681562
; GENERAL INFORMATION:
; APPLICANT: SOBOL, ROBERT E
; APPLICANT: GAGE, FRED H
; APPLICANT: ROYSTON IVOR
; APPLICANT: FRIEDMAN, THEODORE
; TITLE OF INVENTION: LYMPHOKINE GENE THERAPY OF CANCER
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 91211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/352,990
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/863,641
; FILING DATE: 19920403
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-SD 9295
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-9001
; TELEFAX: 619-535-8949
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6365 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: complement (1..6365)

OTHER INFORMATION: /note= "Complementary strand of
FEATURE: PLXSN-RI-1L2"
NAME/KEY: CDS
LOCATION: 2557..3351
US-08-352-990-1

Query Match 28.7%; Score 1486; DB 2; Length 6365;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	3692	GATCTGATCAAGAGACAGATGAGATCGTTTGGCATGATTAAGCAAGATGATGACG	3751
Db	2522	GATCTGATCAAGAGACAGATGAGATCGTTTGGCATGATTAAGCAAGATGATGACG	2581
Qy	3752	CAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTGGCTATGACTGGGCACAAACAGCAA	3811
Db	2582	CAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTGGCTATGACTGGGCACAAACAGCAA	2641
Qy	3812	TGGGCTGCTGATGCGCCGCTGTTCCGGCTGTGACGCGAGGGGCGCCGCTTCTTTTG	3871
Db	2642	TGGGCTGCTGATGCGCCGCTGTTCCGGCTGTGACGCGAGGGGCGCCGCTTCTTTTG	2701
Qy	3872	TCAGACCGACTGTCCGGTCCCTGATGATGAAGTGCAGAGAGGCGCGGCTATCGT	3931
Db	2702	TCAGACCGACTGTCCGGTCCCTGATGATGAAGTGCAGAGAGGCGCGGCTATCGT	2761
Qy	3932	GGCTGGCCACAGCGGGCGCTTCTTGGCAGAGTGTGCTGAGGTTGTCACTGAAGCGGAA	3991
Db	2762	GGCTGGCCACAGCGGGCGCTTCTTGGCAGAGTGTGCTGAGGTTGTCACTGAAGCGGAA	2821
Qy	3992	GGGACTGGCTGATTTGGGGGAGAGTGGCGGGGAGAGTCTCTGTCTCACTTGTCTC	4051
Db	2822	GGGACTGGCTGATTTGGGGGAGAGTGGCGGGGAGAGTCTCTGTCTCACTTGTCTC	2881
Qy	4052	CTGCGGAGAAAGTATCCATCATGCTGATGATCAATGGCGGCTGATGAGCTTGTATCGG	4111
Db	2882	CTGCGGAGAAAGTATCCATCATGCTGATGATCAATGGCGGCTGATGAGCTTGTATCGG	2941
Qy	4112	CTACTCGCCATTTCAGCACCAGCGAAATCATCGCATCGAGCGAGCATCTCGATGG	4171
Db	2942	CTACTCGCCATTTCAGCACCAGCGAAATCATCGCATCGAGCGAGCATCTCGATGG	3001
Qy	4172	AAGCGGCTTGTTCGATCAGGATGATCTGGACGAGAGCATCAGGGGCTCGCGCAGCG	4231
Db	3002	AAGCGGCTTGTTCGATCAGGATGATCTGGACGAGAGCATCAGGGGCTCGCGCAGCG	3061
Qy	4232	AAGTGTTCGCGAGGCTCAAGCGGCGCATGCGCGAGAGATCTCGTGTGACCCATG	4291
Db	3062	AAGTGTTCGCGAGGCTCAAGCGGCGCATGCGCGAGAGATCTCGTGTGACCCATG	3121
Qy	4292	GCGATGCTGCTTCCGGAATCATGTTGGAAATGCGCGCTTCTTGGATTCATCGACT	4351
Db	3122	GCGATGCTGCTTCCGGAATCATGTTGGAAATGCGCGCTTCTTGGATTCATCGACT	3181
Qy	4352	GTGGCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGCTGATTTG	4411
Db	3182	GTGGCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGCTGATTTG	3241
Qy	4412	CTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCGTCTTACGGTATCGCGCTC	4471
Db	3242	CTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCGTCTTACGGTATCGCGCTC	3301
Qy	4472	CCGATTCGAGCGCATCGCCTTCTATCGCTTCTTACGAGTCTTCTGAGCGGACTCT	4531
Db	3302	CCGATTCGAGCGCATCGCCTTCTATCGCTTCTTACGAGTCTTCTGAGCGGACTCT	3361
Qy	4532	GGGTTTCGATAAAATAAAGATTTTATTAGTCTCCAGAAAAGGGGGAATGAAGACC	4591
Db	3362	GGGTTTCGATAAAATAAAGATTTTATTAGTCTCCAGAAAAGGGGGAATGAAGACC	3421
Qy	4592	CCACTGTAGGTTGGCGAGCTAGCTTAAGTAACGCCATTTTTCAGGCGATGGAATAA	4651
Db			

Db	3422	CCACTGTAGGTTGGCGAAGCTAGCTTAAGTAACGCCATTTTTCAGGCGATGGAATAA	3481
Qy	4652	CATAACTGAGAAATAGAGAGATTTAGATCAAGGTTCAGGAAACAGATGGAACAGCTGAATATG	4711
Db	3482	CATAACTGAGAAATAGAGAGATTTAGATCAAGGTTCAGGAAACAGATGGAACAGCTGAATATG	3541
Qy	4712	GGCCAAACAGGATATCTGTGTAGCAGATTTCTCCCGGCTCAGGCGCAAGAACAGATG	4771
Db	3542	GGCCAAACAGGATATCTGTGTAGCAGATTTCTCCCGGCTCAGGCGCAAGAACAGATG	3601
Qy	4772	GAAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAGCAGATTTCTCCCGGCTCAG	4831
Db	3602	GAAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAGCAGATTTCTCCCGGCTCAG	3661
Qy	4832	GGCCAAACAGATGCTCCCGAGATGCGGTCCAGCCCTCAGCAGTTCTTAGAGAACATC	4891
Db	3662	GGCCAAACAGATGCTCCCGAGATGCGGTCCAGCCCTCAGCAGTTCTTAGAGAACATC	3721
Qy	4892	AGATGTTTCCAGGGTCCCGAACAGGACCTGAAATGACCCCTGCTTATTTGAACCTAACCA	4951
Db	3722	AGATGTTTCCAGGGTCCCGAACAGGACCTGAAATGACCCCTGCTTATTTGAACCTAACCA	3781
Qy	4952	ATCAGTTTCGCTTCTCGCTTCTGTTTCGCGCTTCTGCTCCCGAGCTCAATAAAGAGGCC	5011
Db	3782	ATCAGTTTCGCTTCTCGCTTCTGTTTCGCGCTTCTGCTCCCGAGCTCAATAAAGAGGCC	3841
Qy	5012	CACAAACCCCTCACTCGGGGCGCAGTCTCCGATGATGATCGCCGGGTACCCGGT	5071
Db	3842	CACAAACCCCTCACTCGGGGCGCAGTCTCCGATGATGATCGCCGGGTACCCGGT	3901
Qy	5072	ATCCAAATAAACCTTCTGCGAGTTGCGATCCGACTTGTGCTCTCGTGTCTTGGAGGGT	5131
Db	3902	ATCCAAATAAACCTTCTGCGAGTTGCGATCCGACTTGTGCTCTCGTGTCTTGGAGGGT	3961
Qy	5132	CTCCTCTGAGTGTGATGACTACCCGTCAGCGGGGCTCTTTTCAATTTGG	5177
Db	3962	CTCCTCTGAGTGTGATGACTACCCGTCAGCGGGGCTCTTTTCAATTTGG	4007

RESULT 6

US-08-531B-3
; Sequence 3, Application US/08786531B
; Patent No. 6541197
; GENERAL INFORMATION:
; APPLICANT: Link, Charles J.
; APPLICANT: Levy, John P.
; APPLICANT: Wang, Suming
; APPLICANT: Seragina, Tatiana
; TITLE OF INVENTION: Vehicles for Stable Transfer of Green
; TITLE OF INVENTION: Fluorescent Protein Gene and Methods of Use for Same
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
; STREET: 801 Grand Suite 3200
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/786,531B
; APPLICATION NUMBER: US/08/786,531B
; FILING DATE: 21-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: US 60/010371
; APPLICATION NUMBER: US 60/010371
; FILING DATE: 22-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Nebel, Heidi S.
; REGISTRATION NUMBER: 37,719

REFERENCE/DOCKET NUMBER: hqtri
TELECOMMUNICATION INFORMATION:
TELEPHONE: 515-288-3667
TELEFAX: 515-288-1338
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6620 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-786-5318-3

Query Match 28.7%; Score 1486; DB 3; Length 6620;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	3692	GATCTGATCAAGACAGGATGAGGATCGTTTCGCATGATTGACACAGATGGATTGCACG	3751
Db	2777	GATCTGATCAAGACAGGATGAGGATCGTTTCGCATGATTGACACAGATGGATTGCACG	2836
Qy	3752	CAGGTTCTCCGGCGCTTGGGTGAGAGGCTATTTCGGCTATGACTGGGCAACACAGACAA	3811
Db	2837	CAGGTTCTCCGGCGCTTGGGTGAGAGGCTATTTCGGCTATGACTGGGCAACACAGACAA	2896
Qy	3812	TCGGCTGCTGTATGTCGGCCGTGTTCCGGCTGTTCAGGCGAGGGCGCCGGTCTTTTTCG	3871
Db	2897	TCGGCTGCTGTATGTCGGCCGTGTTCCGGCTGTTCAGGCGAGGGCGCCGGTCTTTTTCG	2956
Qy	3872	TCAGACCGACTGTTCGGTCCCTGATGATGAATGCACTGACGAGGACGAGCGGGTATTCGT	3931
Db	2957	TCAGACCGACTGTTCGGTCCCTGATGATGAATGCACTGACGAGGACGAGCGGGTATTCGT	3016
Qy	3932	GGCTGGCCACACAGGGCGTCTTCGGCAGCTGTCTGACGTGTGCTCACTGAAGCGGAA	3991
Db	3017	GGCTGGCCACACAGGGCGTCTTCGGCAGCTGTCTGACGTGTGCTCACTGAAGCGGAA	3076
Qy	3992	GGGACTGGCTGTATTTGGGCGAAGTGGCGGGCAGGATCTCTGTGATCTCACCTTGCTC	4051
Db	3077	GGGACTGGCTGTATTTGGGCGAAGTGGCGGGCAGGATCTCTGTGATCTCACCTTGCTC	3136
Qy	4052	CTGGCGAAGATATCATATGCTGATGATGATGATGATGATGATGATGATGATGATGATG	4111
Db	3137	CTGGCGAAGATATCATATGCTGATGATGATGATGATGATGATGATGATGATGATGATG	3196
Qy	4112	CTACTCGCCATTTCGACCAACAGCAATCATGCTGACGAGCAGCTACTCGGATGG	4171
Db	3197	CTACTCGCCATTTCGACCAACAGCAATCATGCTGACGAGCAGCTACTCGGATGG	3256
Qy	4172	AAGCCGGTCTTTGCTGATCAGGATGATCTGGACGAGCAGCATCAGGGGCTCGCCAGCCG	4231
Db	3257	AAGCCGGTCTTTGCTGATCAGGATGATCTGGACGAGCAGCATCAGGGGCTCGCCAGCCG	3316
Qy	4232	AAGTGTTCGACAGGCTCAAGGGCGCATCGCCGAGCGGAGGATCTGCTGTGACCCATG	4291
Db	3317	AAGTGTTCGACAGGCTCAAGGGCGCATCGCCGAGCGGAGGATCTGCTGTGACCCATG	3376
Qy	4292	CGGATGCTGCTTTCGGAATATCATGTTGGAATGAGCGGCTTTTCTGGAATTCAGCT	4351
Db	3377	CGGATGCTGCTTTCGGAATATCATGTTGGAATGAGCGGCTTTTCTGGAATTCAGCT	3436
Qy	4352	GTGGCCGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGGTGATATG	4411
Db	3437	GTGGCCGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGGTGATATG	3496
Qy	4412	CTGAAGAGCTTGGCGGCAATGGGCTGACCGCTTCTGCTGCTTTACGGTATCGCGCTC	4471
Db	3497	CTGAAGAGCTTGGCGGCAATGGGCTGACCGCTTCTGCTGCTTTACGGTATCGCGCTC	3556
Qy	4472	CCGATTCGAGGATCGCCCTTCTATCGCTTCTGAGAGTCTTCTGAGCGGACTCT	4531

Db	3557	CCGATTCGAGCGCATCGCCTTCTATCGCCTCTTGTAGCAGTTCCTTGTAGCGGACTCT	3616
Qy	4532	GGGGTTCGATAAATAAAGATTTTATTTAGTCTCCAGAAAGGGGGGATGAAGACC	4591
Db	3617	GGGGTTCGATAAATAAAGATTTTATTTAGTCTCCAGAAAGGGGGGATGAAGACC	3676
Qy	4592	CCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAAAGCCATTTTGAAGGCAATGAAAAATA	4651
Db	3677	CCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAAAGCCATTTTGAAGGCAATGAAAAATA	3736
Qy	4652	CATAACTGAGATAGAGAGTTCAGATCAAGGTCAGAAACAGATGGAACAGCTGAATATG	4711
Db	3737	CATAACTGAGATAGAGAGTTCAGATCAAGGTCAGAAACAGATGGAACAGCTGAATATG	3796
Qy	4712	GGCCAAACAGGATATCTGTGTAAGCAGTTCTTCCGCCCGCTCAGGGCCAAAGAACAGATG	4771
Db	3797	GGCCAAACAGGATATCTGTGTAAGCAGTTCTTCCGCCCGCTCAGGGCCAAAGAACAGATG	3856
Qy	4772	GAACAGCTGAATATGGCCAAACAGGATATCTGTGTAAGCAGTTCTTCCGCCCGCTCAG	4831
Db	3857	GAACAGCTGAATATGGCCAAACAGGATATCTGTGTAAGCAGTTCTTCCGCCCGCTCAG	3916
Qy	4832	GGCCAAAGACAGATGTTCCCGCAGATGCGGTCCAGCCTCAGCAGTTCCTAGAGAACCATC	4891
Db	3917	GGCCAAAGACAGATGTTCCCGCAGATGCGGTCCAGCCTCAGCAGTTCCTAGAGAACCATC	3976
Qy	4892	AGATGTTTCAGGGTGGCCCAAGGACCTGAATGACCTGTGCTTATTTTGAACCTAACCA	4951
Db	3977	AGATGTTTCAGGGTGGCCCAAGGACCTGAATGACCTGTGCTTATTTTGAACCTAACCA	4036
Qy	4952	ATCAGTTTCGCTTCTCGCTTCTGTTGCGCGCTTCTGCTCCCGAGCTCAATAAAGAGCC	5011
Db	4037	ATCAGTTTCGCTTCTCGCTTCTGTTGCGCGCTTCTGCTCCCGAGCTCAATAAAGAGCC	4096
Qy	5012	CACAAACCCCTCACTCGGGGCGCCAGTCTCCGATGACTGAGTCCGCCGGGTACCCGTGT	5071
Db	4097	CACAAACCCCTCACTCGGGGCGCCAGTCTCCGATGACTGAGTCCGCCGGGTACCCGTGT	4156
Qy	5072	ATCCAAATAAACCTCTTGGAGTTGCTATCCGACTTGTGCTGCTGCTTCTTGGAGGGT	5131
Db	4157	ATCCAAATAAACCTCTTGGAGTTGCTATCCGACTTGTGCTGCTGCTTCTTGGAGGGT	4216
Qy	5132	CTCTCTGAGTGTGATGACTACCGTCAAGCGGGGTCTTTTCATTTGG	5177
Db	4217	CTCTCTGAGTGTGATGACTACCGTCAAGCGGGGTCTTTTCATTTGG	4262

RESULT 7
US-07-753-5208-3
Sequence 3, Application US/07753520B
Patent No. 5352595
GENERAL INFORMATION:
APPLICANT: Tapscott, J.; Weintraub, H. M.; Palmer, T. D.
TITLE OF INVENTION: "MYOD REGULATORY REGION"
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
STREET: 2800 Pacific First Center, 1420 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette-5.25 inch, 1.2Mb storage
COMPUTER: IBM PC/386 Compatible
OPERATING SYSTEM: MS-DOS 4.01
SOFTWARE: Word for Windows-t
CURRENT APPLICATION DATA: US/07753,520B
APPLICATION NUMBER: 19910903
FILING DATE: 19910903
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: none

; FILING DATE: none
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sundemo, John S.
 ; REGISTRATION NUMBER: 34,446
 ; REFERENCE/DOCKET NUMBER: PHCR-1-5789
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 1-206-682-8100; 1-206-224-0727 (direct)
 ; TELEFAX: 1-206-224-0779
 ; TELEX: 4938023
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 9115 base pairs
 ; TYPE: NUCLEIC ACID
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: Other; plasmid DNA
 ; DESCRIPTION: PLHMDN-53: 5'LTR (position 1-1159); Y+ (position
 ; DESCRIPTION: 1159-1640); HisD (position 1641-3007); Myo-D 53L.4 ApaI fragment
 ; DESCRIPTION: (position 3008-5248); driving neo (position 5249-6117); with a
 ; DESCRIPTION: 3'LTR (position 6118-6823) coupled to a pBR322 plasmid (position
 ; DESCRIPTION: 6824-9115); Figures 7A-7D.
 ; US-07-753-520B-3

5932	GTGGCGGCTGGGTGTGGGGACCGGCTATCAGACACATAGGTTGGCTACCGGTGATATTG	5991
4412	CTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCTGTCGTTTACGGTATCGCGCTC	4471
5992	CTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCTGTCGTTTACGGTATCGCGCTC	6051
4472	CCGATTGCGACGCATCGCCTTCTATCGCCTTCTTGAAGAGTTCTTCTGAGCGGACTCT	4531
6052	CCGATTGCGACGCATCGCCTTCTATCGCCTTCTTGAAGAGTTCTTCTGAGCGGACTCT	6111
4532	GGGGTTCGATAAAATAAAGATTTTATTTAGTCTCCAGAAAAGGGGGNATGAAGACC	4591
6112	GGGGTTCGATAAAATAAAGATTTTATTTAGTCTCCAGAAAAGGGGGNATGAAGACC	6171
4592	CCACCTGTAGTTTGCACAGCTAGCTTAAGTACGCCATTTTCTCAAGGCAATGAAAATA	4651
6172	CCACCTGTAGTTTGCACAGCTAGCTTAAGTACGCCATTTTCTCAAGGCAATGAAAATA	6231
4652	CATAACTGAGATAGAGAGAGTTCCAGATCAAGGTTCAGGACAGATGGAACAGCTGAATATG	4711
6232	CATAACTGAGATAGAGAGAGTTCCAGATCAAGGTTCAGGACAGATGGAACAGCTGAATATG	6291
4712	GGCCAAACAGGATATCTGTGTGAAGCAGTTCTCTGCCCGGCTCAGGGCCCAAGAACAGATG	4771
6292	GGCCAAACAGGATATCTGTGTGAAGCAGTTCTCTGCCCGGCTCAGGGCCCAAGAACAGATG	6351
4772	GAACAGCTGAATATGGGCCAAACAGGATATCTGTGTGAAGCAGTTCTCTGCCCGGCTCAG	4831
6352	GAACAGCTGAATATGGGCCAAACAGGATATCTGTGTGAAGCAGTTCTCTGCCCGGCTCAG	6411
4832	GGCCAGAACAGATGTTGCCCGCAGATCGGGTCCAGCCCTCAGCAGTTTCTAGAGAACCATC	4891
6412	GGCCAGAACAGATGTTGCCCGCAGATCGGGTCCAGCCCTCAGCAGTTTCTAGAGAACCATC	6471
4892	AGATGTTCCAGGGTCCCGCAGGACTGAATGACCTGTGCTTATTTTGAACCTAACCA	4951
6472	AGATGTTCCAGGGTCCCGCAGGACTGAATGACCTGTGCTTATTTTGAACCTAACCA	6531
4952	ATCAGTTTCGGTCTCTCGTCTTGTTCGGCGCTTCTGCTCCCGAGCTCAATAAAGAGCC	5011
6532	ATCAGTTTCGGTCTCTCGTCTTGTTCGGCGCTTCTGCTCCCGAGCTCAATAAAGAGCC	6591
5012	CACAACCCCTCACTCGGGGGCGCAGTCTCTCCGATTGACTGAGTCGCCCGGGTACCCGTGT	5071
6592	CACAACCCCTCACTCGGGGGCGCAGTCTCTCCGATTGACTGAGTCGCCCGGGTACCCGTGT	6651
5072	ATCCAAATAAACCTCTTTCGAGTTGCATTCGACTTGTGGTCTCGCTGTTCCTTGGGAGGGT	5131
6652	ATCCAAATAAACCTCTTTCGAGTTGCATTCGACTTGTGGTCTCGCTGTTCCTTGGGAGGGT	6711
5132	CTCCTCTGAGTGATGACTACCCGTCAGCGGGGGTCTTTCATTTTGG	5177
6712	CTCCTCTGAGTGATGACTACCCGTCAGCGGGGGTCTTTCATTTTGG	6757

```

RESULT 8
US-08-654-737B-3
; Sequence 3, Application US/08654737B
; Patent No. 6274136
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; APPLICANT: Weiner, Leslie P.
; APPLICANT: McMillan, Minnie
; TITLE OF INVENTION: CONSTRUCTION AND USE OF GENES ENCODING
; TITLE OF INVENTION: PATHOGENIC EPITOPES FOR TREATMENT OF AUTOIMMUNE DISEASE
; FILE REFERENCE: 13761-703-00 US
; CURRENT APPLICATION NUMBER: US/08/654,737B
; CURRENT FILING DATE: 1996-05-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5865
; TYPE: DNA

```


ORGANISM: Artificial Sequence

FEATURE: Retroviral vector derived from Moloney Murine

OTHER INFORMATION: Leukemia Virus

US-08-654-737B-3

Query Match 28.1%; Score 1455.2; DB 3; Length 5865;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 1488; Conservative 0; Mismatches 3; Indels 21; Gaps 1;

QY 3687 ATCAAGATCTGATCAAGACAGAGGATGAGGATCGTTTCGCGATGATGAAACAGATGGAATT 3746
DB 1852 ATTCCGATCTGATCAAGACAGAGGATGAGGATCGTTTCGCGATGATGAAACAGATGGAATT 1911
QY 3747 GCACGACGATCTCCGGCCGCTTGGGTGAGAGGCTATTCCGCTATGACTGCGGCACACACA 3806
DB 1912 GCACGACGATCTCCGGCCGCTTGGGTGAGAGGCTATTCCGCTATGACTGCGGCACACACA 1971
QY 3807 GACAAATCGGCTGCTCTGATGCCCGCTGTTCCGGCTGTGACGCGAGGGCGCCCGGTTCT 3866
DB 1972 GACAAATCGGCTGCTCTGATGCCCGCTGTTCCGGCTGTGACGCGAGGGCGCCCGGTTCT 2031
QY 3867 TTTTGTCAAGACCGACTGTCGGTGCCTGAAATGAATGACGACGAGGCGCGGCT 3926
DB 2032 TTTTGTCAAGACCGACTGTCGGTGCCTGAAATGAATGACGACGAGGCGCGGCT 2091
QY 3927 ATCTGGGCTGGCCAGCGGGGCTTCTTGGCGAGCTGTGCTCGACGTTGTCACCTGAAGC 3986
DB 2092 ATCTGGGCTGGCCAGCGGGGCTTCTTGGCGAGCTGTGCTCGACGTTGTCACCTGAAGC 2151
QY 3987 GGGAGGAGCTGGCTGCTATTGGGCGAAGTGC CGGGGAGGATCTCTGTGTCATCTCACCT 4046
DB 2152 GGGAGGAGCTGGCTGCTATTGGGCGAAGTGC CGGGGAGGATCTCTGTGTCATCTCACCT 2211
QY 4047 TGCTCTCCGAGAAAGTATCATATGCTGATGCAATGCGGCTGATGCAATGCGGCTGATGCA 4106
DB 2212 TGCTCTCCGAGAAAGTATCATATGCTGATGCAATGCGGCTGATGCAATGCGGCTGATGCA 2271
QY 4107 TCCGGCTACCTGCCATTCGACCCAGCCAGCAGCAATCGCATCGAGCGACACTACTCG 4166
DB 2272 TCCGGCTACCTGCCATTCGACCCAGCCAGCAGCAATCGCATCGAGCGACACTACTCG 2331
QY 4167 GATGGAAGCCGGTCTTGTGATCAGGATGATCTGGAAGAGAGCATCAGGGGCTCGCGCC 4226
DB 2332 GATGGAAGCCGGTCTTGTGATCAGGATGATCTGGAAGAGAGCATCAGGGGCTCGCGCC 2391
QY 4227 AGCGGAATGTTGCGCAGGCTCAAGGCGCGCATGCCGAGCGGAGGATCTCGTCTGAC 4286
DB 2392 AGCGGAATGTTGCGCAGGCTCAAGGCGCGCATGCCGAGCGGAGGATCTCGTCTGAC 2451
QY 4287 CCATGGCGATCGCTGCTTGCAGATATCATGTTGGAATGCGGCTTCTTCTGATTCAT 4346
DB 2452 CCATGGCGATCGCTGCTTGCAGATATCATGTTGGAATGCGGCTTCTTCTGATTCAT 2511
QY 4347 CGACTGTGGCGGCTGTTGGCGGAGCGCTATCAGGACATAGCTTGGCTACCCGTGA 4406
DB 2512 CGACTGTGGCGGCTGTTGGCGGAGCGCTATCAGGACATAGCTTGGCTACCCGTGA 2571
QY 4407 TATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCGTCTTTACGGTATCGC 4466
DB 2572 TATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCGTCTTTACGGTATCGC 2631
QY 4467 CGCTCCCGATTCGACGCGCATCGCTTCTATCGGCTTCTTGAACGAGTCTTCTGAGCGG 4526
DB 2632 CGCTCCCGATTCGACGCGCATCGCTTCTATCGGCTTCTTGAACGAGTCTTCTGAGCGG 2691
QY 4527 ACTCTGGGGT-----TCGATAAAATAAAGATTTTATTAGTCT 4565
DB 2692 ACTCTGGGGTTCGTCGAGAGCTTGGGCGCATCGATAAAATAAAGATTTTATTAGTCT 2751
QY 4566 CCAGAAAAGGGGGAATGAAAGACCCCACTGTAGGTTTGGCAAGCTTAGCTTAAGTAAC 4625
DB 2752 CCAGAAAAGGGGGAATGAAAGACCCCACTGTAGGTTTGGCAAGCTTAGCTTAAGTAAC 2811

QY 4626 GCCATTTTGAAGGCATGGAAAAATACATAACTGAGAATAGAGAAGTTTCAGATCAAGGTC 4685
DB 2812 GCCATTTTGAAGGCATGGAAAAATACATAACTGAGAATAGAGAAGTTTCAGATCAAGGTC 2871
QY 4686 AGGAACAGATGGAAACAGCTGAATATGGGCCCAACAGAGATATCTGTGTAAAGCAGTCTCTG 4745
DB 2872 AGGAACAGATGGAAACAGCTGAATATGGGCCCAACAGAGATATCTGTGTAAAGCAGTCTCTG 2931
QY 4746 CCCCCTCAGGGCCCAAGAACAGATGGAACAGCTGAATATGGGCCCAACAGAGATATCTGT 4805
DB 2932 CCCCCTCAGGGCCCAAGAACAGATGGAACAGCTGAATATGGGCCCAACAGAGATATCTGT 2991
QY 4806 GGTAAAGCAGTCTCTGCCCGGCTCAGGGCCCAAGAACAGATGGAACAGCTGAATATGGGCCCAAG 4865
DB 2992 GGTAAAGCAGTCTCTGCCCGGCTCAGGGCCCAAGAACAGATGGAACAGCTGAATATGGGCCCAAG 3051
QY 4866 CCTCAGCAGTCTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCCAAGGACCTGAAATG 4925
DB 3052 CCTCAGCAGTCTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCCAAGGACCTGAAATG 3111
QY 4926 ACCCTGTGCTTATTGAACTAACCAATCAGTTCGCTTCTCGCTTCTGTTCGCGCGCTTC 4985
DB 3112 ACCCTGTGCTTATTGAACTAACCAATCAGTTCGCTTCTCGCTTCTGTTCGCGCGCTTC 3171
QY 4986 TGCTCCCGGAGCTCAATAAAGAGAGCCCAACACCTCCTCAGTTCGGGCGCCAGTCTCCGAT 5045
DB 3172 TGCTCCCGGAGCTCAATAAAGAGAGCCCAACACCTCCTCAGTTCGGGCGCCAGTCTCCGAT 3231
QY 5046 TGACTGAGTCGCGGGTACCCGCTGATCAATAAACCCTCTTTCAGTTCGATCCGACTT 5105
DB 3232 TGACTGAGTCGCGGGTACCCGCTGATCAATAAACCCTCTTTCAGTTCGATCCGACTT 3291
QY 5106 GTGCTCTGCTGTTCTTGGGAGGCTCTCTCTGAGTGATGACTACCGTACGCGGGG 5165
DB 3292 GTGCTCTGCTGTTCTTGGGAGGCTCTCTCTGAGTGATGACTACCGTACGCGGGG 3351
QY 5166 TCTTTTCATTGG 5177
DB 3352 TCTTTTCATTGG 3363

RESULT 9

US-08-462-859A-6
; Sequence 6, Application US/08462859A
; Patent No. 5652092
; GENERAL INFORMATION:
; APPLICANT: Jacobsen, J. S.
; APPLICANT: Vitek, M. P.
; TITLE OF INVENTION: No. 5652092el Amyloid Precursor and Method of
; TITLE OF INVENTION: Using Same to Access Agents Which Down-Regulate Formation
; TITLE OF INVENTION: of B-Amyloid Peptide
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Cyanamid Company
; STREET: One Cyanamid Plaza
; CITY: Wayne
; STATE: New Jersey
; COUNTRY: United States
; ZIP: 07470-8426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,859A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Barnhard, Elizabeth M.
; REGISTRATION NUMBER: 31,088
; REFERENCE/DOCKET NUMBER: 31,844-04

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Barnhard, Elizabeth M.
REGISTRATION NUMBER: 31,088
REFERENCE/DOCKET NUMBER: 31,844-04
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201)831-3246
TELEFAX: (201)831-3305
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 8591 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2393..3853
US-08-462-859A-8

Query Match 28.0%; Score 1448.4; DB 2; Length 8591;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1474; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

QY	3062	TCCCTCAGGATATAGTAGTTTCGCTTTGATAGGAGGGGAAATGTAGTCTTATGCA	3121
DB	4931	TCCCTCAGGATATAGTAGTTTCGCTTTGATAGGAGGGGAAATGTAGTCTTATGCA	4990
QY	3122	ATATCTTTGTAGTCTTGCAACATGGTAAACGATAGTTAGCAACATGCTTACAAAGGAG	3181
DB	4991	ATACACTTTGTAGTCTTGCAACATGGTAAACGATAGTTAGCAACATGCTTACAAAGGAG	5050
QY	3182	AAAAAGCCGTCATGCGGATGGTGAAGTGGTGGTACGATCGTGCCTTATTAGGA	3241
DB	5051	AAAAAGCACCGTCATGCGGATGGTGAAGTGGTGGTACGATCGTGCCTTATTAGGA	5110
QY	3242	AGGCAACAGACGGGTCTGACATGATTCGACGAAACCACTGAATTCGCGATTTCAGAGAT	3300
DB	5111	AGGCAACAGACGGGTCTGACATGATTCGACGAAACCACTGAATTCGCGATTTCAGAGAT	5170
QY	3301	ATTGTATTAAAGTGCCTAGCTCGATACAGCAAAACGCCATTTGACCATTTCCACCATTTGGT	3360
DB	5171	ATTGTATTAAAGTGCCTAGCTCGATACAAATAAACGCCATTTGACCATTTCCACCATTTGGT	5230
QY	3361	GTGCACCTCCAGCTTACGCTGCGGCAAGCACTCAGGGCCCAAGGGCTGCTAAAGGAG	3420
DB	5231	GTGCACCTCCAGCTTACGCTGCGGCAAGCACTCAGGGCCCAAGGGCTGCTAAAGGAG	5290
QY	3421	CGGAACAGTAAAGACAGTCCGAGAAACGGGTGCTGACCCCGATGAATGTACAGCTAC	3480
DB	5291	CGGAACAGTAAAGACAGTCCGAGAAACGGGTGCTGACCCCGATGAATGTACAGCTAC	5350
QY	3481	TGGGCTATCTGGACAGGGAACACGCAAGCCAAAGAGAAAGAGGTAGCTTGCAAGTGGG	3540
DB	5351	TGGGCTATCTGGACAGGGAACACGCAAGCCAAAGAGAAAGAGGTAGCTTGCAAGTGGG	5410
QY	3541	CTTACATGGCGATAGCTAGCTGGCGGTTTATGGAAGCAAGCAAGCAAGCAAGCAAGTGGCA	3600
DB	5411	CTTACATGGCGATAGCTAGCTGGCGGTTTATGGAAGCAAGCAAGCAAGCAAGTGGCA	5470
QY	3601	GCTGGGCGGCTCTGGTAAGTTGGGAAGCCCTGCAAAAGTAAACTGATGGCTTTCTTG	3660
DB	5471	GCTGGGCGGCTCTGGTAAGTTGGGAAGCCCTGCAAAAGTAAACTGATGGCTTTCTTG	5530
QY	3661	CCGCCAAGGATCTGATGGCGAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCG	3720
DB	5531	CCGCCAAGGATCTGATGGCGAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCG	5590
QY	3721	TTTCGATGATGAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	3780
DB	5591	TTTCGATGATGAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	5650
QY	3781	CTATTGGCTATGACTGGGCACACAGACAATCGGCTCTGATGCGCGCTGTTCCGG	3840

DB	5651	CTATTGGCTATGACTGGGCACACAGACAATCGGCTGCTCTGATGCGCGCTGTTCCGG	5710
QY	3841	CTGTACGCGAGGGGCCCGGTTCTTTTGTCAAGACCGACCTGTGCGGTGCTTGAAT	3900
DB	5711	CTGTACGCGAGGGGCCCGGTTCTTTTGTCAAGACCGACCTGTGCGGTGCTTGAAT	5770
QY	3901	GAATGACAGACGAGGCGCGCTATCGTGGCTGCGCACAGCGGCGTTCCTTGGCA	3960
DB	5771	GAATGACAGACGAGGCGCGCTATCGTGGCTGCGCACAGCGGCGTTCCTTGGCA	5830
QY	3961	GCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGGCTGTATTGGCGAAGTGC	4020
DB	5831	GCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGGCTGTATTGGCGAAGTGC	5890
QY	4021	GGGAGGATCTCTGTCTCATCTCACCTTGTCTCTGCGAGAAAGTATCATCATGGCTG	4080
DB	5891	GGGAGGATCTCTGTCTCATCTCACCTTGTCTCTGCGAGAAAGTATCATCATGGCTG	5950
QY	4081	GCAATGCGCGGCTGCATACGCTTGATCGGCTACCTGCGCATTCGACCAACCAAGCGAA	4140
DB	5951	GCAATGCGCGGCTGCATACGCTTGATCGGCTACCTGCGCATTCGACCAACCAAGCGAA	6010
QY	4141	CATCGCATCGAGGACGACGCTACTCGGATGGAAGCGGCTCTTGTGATCAGGATGATG	4200
DB	6011	CATCGCATCG- GCGAGCACGCTACTCGGATGGAAGCGGCTCTTGTGATCAGGATGATG	6069
QY	4201	GACGAAGACATCAGGGGCTCGGCGCAGCGCACTGTTGCCAGGCTCAAGGCGCGCATG	4260
DB	6070	GACGAAGACATCAGGGGCTCGGCGCAGCGCACTGTTGCCAGGCTCAAGGCGCGCATG	6129
QY	4261	CCGAGCGGAGGATCTGCTGACCATGCGGATGCGCTGCTTGCAGAAATATCATGGTG	4320
DB	6130	CCGAGCGGAGGATCTGCTGACCATGCGGATGCGCTGCTTGCAGAAATATCATGGTG	6189
QY	4321	GAATAAGCGCGCTTTCTGGAATTCATCGACTGTGGCGGCTGGGTGTGGCGGACCGCTAT	4380
DB	6190	GAATAAGCGCGCTTTCTGGAATTCATCGACTGTGGCGGCTGGGTGTGGCGGACCGCTAT	6249
QY	4381	CAGCAATAGCGTTGGTACCGCTGATATTCGTAAGAGCTTGGCGGCGAATGGGCTGAC	4440
DB	6250	CAGCAATAGCGTTGGTACCGCTGATATTCGTAAGAGCTTGGCGGCGAATGGGCTGAC	6309
QY	4441	CGCTTCTCTGCTTTTACGCTATCGCGCTCCGATTCGAGCGCATCGCTTCTATCGC	4500
DB	6310	CGCTTCTCTGCTTTTACGCTATCGCGCTCCGATTCGAGCGCATCGCTTCTATCGC	6369
QY	4501	CTTCTTGACGAGTCTTCTGAGCGGACTCTGGGGTTTCGATA	4542
DB	6370	CTTCTTGACGAGTCTTCTGAGCGGACTCTGGGGTTTCGATA	6411

RESULT 11

US-08-123-659A-6
Sequence 6, Application US/08123659A
Patent No. 5656477
GENERAL INFORMATION:
APPLICANT: Jacobsen, J. S.
APPLICANT: Vitek, M. P.
TITLE OF INVENTION: No. 5656477el Amyloid Precursor and Method of
TITLE OF INVENTION: Using Same to Access Agents Which Down-Regulate Formation
TITLE OF INVENTION: of B-Amyloid Peptide
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Anne Rosenblum
STREET: 163 Delaware Avenue, Suite 212
CITY: Delmar
STATE: New York
COUNTRY: U.S.A.
ZIP: 12054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

; CORRESPONDENCE ADDRESS:
 ; ADDRESS: American Cyanamid Company
 ; STREET: One Campus Drive
 ; CITY: Parsippany
 ; STATE: New Jersey
 ; COUNTRY: United States
 ; ZIP: 07054
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/464,247A
 ; FILING DATE: 05-JUN-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Barnhard, Elizabeth M.
 ; REGISTRATION NUMBER: 31,088
 ; REFERENCE/DOCKET NUMBER: 31,844-03
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-683-2158
 ; TELEFAX: 201-683-4117
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 5991 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: circular
 ; MOLECULE TYPE: CDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 2393..3868
 ; US-08-464-247A-6

Query Match 28.0%; Score 1448.4; DB 2; Length 8591;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1474; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

Qy	3062	TCCCTCAGGATATAGTATGTTTCGCTTTTGCATAGGAGGGGAAATGTAGTCTTATGCA	3121
Db	4931	TCCCTCAGGATATAGTATGTTTCGCTTTTGCATAGGAGGGGAAATGTAGTCTTATGCA	4990
Qy	3122	ATACCTTTGTAGTCTTGCACATGTAACGATGATGACCAACATGCTTAAAGGAGAG	3181
Db	4991	ATACCTTTGTAGTCTTGCACATGTAACGATGATGACCAACATGCTTAAAGGAGAG	5050
Qy	3182	AAAAGCACCGTGCATGCGGATGTTGGAAGTAAGTGTGATGCTGCTTATTAGGA	3241
Db	5051	AAAAGCACCGTGCATGCGGATGTTGGAAGTAAGTGTGATGCTGCTTATTAGGA	5110
Qy	3242	AGGCAACAGACGGGTCTGACATGATTCGACGAACCACTGAATTCGCAATTCAGAGAT	3300
Db	5111	AGGCAACAGACGGGTCTGACATGATTCGACGAACCACTGAATTCGCAATTCAGAGATA	5170
Qy	3301	ATTGTATTAAAGTGTCTAGCTGATACAGCAACGCCATTTGACCAATTCACCAATTGGT	3360
Db	5171	ATTGTATTAAAGTGTCTAGCTGATACCAATTAAGGCCATTTGACCAATTCACCAATTGGT	5230
Qy	3361	GTGCACTTCAGCTTCACTGCTGCGGAGCACTCAGGCGGCAAGGCTGCTTAAGGAAG	3420
Db	5231	GTGCACTTCAGCTTCACTGCTGCGGAGCACTCAGGCGGCAAGGCTGCTTAAGGAAG	5290
Qy	3421	CGGAACAGTGAAGACCGATCCGAGAAACGGTCTGACCCCGGATGAATGTCAGCTAC	3480
Db	5291	CGGAACAGTGAAGACCGATCCGAGAAACGGTCTGACCCCGGATGAATGTCAGCTAC	5350
Qy	3481	TGGGCTATCTGCAAGGGAACCAAGCGCAAGAGAAAGCAGGTAGCTTGCAGTGGG	3540
Db	5351	TGGGCTATCTGCAAGGGAACCAAGCGCAAGAGAAAGCAGGTAGCTTGCAGTGGG	5410
Qy	3541	CTTACATGCGGATAGCTAGCTGCGGGGTTTATGCAAGCAAGCGAACCAGGAATTGCA	3600

Db	5411	CTTACATGCGGATAGCTAGCTGCGGGTTTTATGGACAGCAAGCAACCGGAATTGCCA	5470
Qy	3601	GCTGGGGCGCCTCTGTGTAAGTTGGGAAGCCCTGCAAAAGTAAACTGATGGCTTTCTTG	3660
Db	5471	GCTGGGGCGCCTCTGTGTAAGTTGGGAAGCCCTGCAAAAGTAAACTGATGGCTTTCTTG	5530
Qy	3661	CCGCCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGATCG	3720
Db	5531	CCGCCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGATCG	5590
Qy	3721	TTTCGCATGATTGAACAAGATGGATTGACACAGGTTCTCCGGCGCTTGGGTGGAGAGG	3780
Db	5591	TTTCGCATGATTGAACAAGATGGATTGACACAGGTTCTCCGGCGCTTGGGTGGAGAGG	5650
Qy	3781	CTATTTCGGCTATGATCGGGCAACAAGCAATCGGCTCTGATGCGCGCTGTTCCGG	3840
Db	5651	CTATTTCGGCTATGATCGGGCAACAAGCAATCGGCTCTGATGCGCGCTGTTCCGG	5710
Qy	3841	CTGTACGGCAGGGCGCGGCTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTGAAT	3900
Db	5711	CTGTACGGCAGGGCGCGGCTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTGAAT	5770
Qy	3901	GAACTGCAGGACAGGCGCGCTATCGTGGCTGGCCACGACGGGGCTTCTTGGCGCA	3960
Db	5771	GAACTGCAGGACAGGCGCGCTATCGTGGCTGGCCACGACGGGGCTTCTTGGCGCA	5830
Qy	3961	GCTGTGCTCGACGTTGTCACTGAAGCGGAAAGGAGTGGTGTCTATTGGGCGAAGTCCG	4020
Db	5831	GCTGTGCTCGACGTTGTCACTGAAGCGGAAAGGAGTGGTGTCTATTGGGCGAAGTCCG	5890
Qy	4021	GGGCAGGATCTCTGTCTCATCTCACCTTCTCTCCCGAGAAAGTATCCATCATGGCTGAT	4080
Db	5891	GGGCAGGATCTCTGTCTCATCTCACCTTCTCTCCCGAGAAAGTATCCATCATGGCTGAT	5950
Qy	4081	GCAATGGCGGCTGTCATAGCTTGTATCGGCTACTTGCCTTTCGACCAACCAAGCGAAA	4140
Db	5951	GCAATGGCGGCTGTCATAGCTTGTATCGGCTACTTGCCTTTCGACCAACCAAGCGAAA	6010
Qy	4141	CATCGCATCG-CCGAGCAGCTACTCGATGGAAGCGGCTCTTGTGATCAGGATGATCTG	4200
Db	6011	CATCGCATCG-CCGAGCAGCTACTCGATGGAAGCGGCTCTTGTGATCAGGATGATCTG	6069
Qy	4201	GACGAAGCATCAGGGGCTCGCCGACCGCAACTGTTGCGCAGGCTCAAGCGCGGATG	4260
Db	6070	GACGAAGCATCAGGGGCTCGCCGACCGCAACTGTTGCGCAGGCTCAAGCGCGGATG	6129
Qy	4261	CCCCGCGCAGGATCTGCTGACCCCATGCGGCTGCTTGTGCGCAATATCATGTTG	4320
Db	6130	CCCCGCGCAGGATCTGCTGACCCCATGCGGCTGCTTGTGCGCAATATCATGTTG	6189
Qy	4321	GAATAATGGCGCTTTTCTGATTCATCGACTGTGCGCGGCTGGGTGTGCGGACCGCTAT	4380
Db	6190	GAATAATGGCGCTTTTCTGATTCATCGACTGTGCGCGGCTGGGTGTGCGGACCGCTAT	6249
Qy	4381	CAGGACATAGCGTTGGCTACCCGTTGATTTCTGAAAGAGCTTGGCGCGAATGGGCTGAC	4440
Db	6250	CAGGACATAGCGTTGGCTACCCGTTGATTTCTGAAAGAGCTTGGCGCGAATGGGCTGAC	6309
Qy	4441	CGCTTCCTCGTCTTACGGTATCGCGCTCCCGATTCGACGCGCATCGCTTCTTATCGC	4500
Db	6310	CGCTTCCTCGTCTTACGGTATCGCGCTCCCGATTCGACGCGCATCGCTTCTTATCGC	6369
Qy	4501	CTTCTTGACGAGTTCTTCTGAGCGGAGCTCTGGGGTTCGATA 4542	
Db	6370	CTTCTTGACGAGTTCTTCTGAGCGGAGCTCTGGGGTTCGATA 6411	

RESULT 14
 US-08-464-247A-8
 ; Sequence 8, Application US/08464247A
 ; Patent No. 5693478
 ; GENERAL INFORMATION:
 ; APPLICANT: Jacobsen, J. S.

APPLICANT: Vitek, M. P.
TITLE OF INVENTION: No. 5693478el Amyloid Precursor and Method of
TITLE OF INVENTION: Using Same to Access Agents Which Down-Regulate Formation
TITLE OF INVENTION: of B-Amyloid Peptide
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESS: American Cyanamid Company
STREET: One Campus Drive
CITY: Parsippany
STATE: New Jersey
COUNTRY: United States
ZIP: 07054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,247A
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Barnhard, Elizabeth M.
REGISTRATION NUMBER: 31,088
REFERENCE/DOCKET NUMBER: 31,844-03
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-683-2158
TELEFAX: 201-683-4117
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 8591 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2393..3853
US-08-464-247A-8

Query Match 28.0%; Score 1448.4; DB 2; Length 8591;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1474; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

QY	3062	TCCCTCAGGATATAGTATGTTTCGTTTTCATAGGAGGGGAAATAGTCTTATGCA	3121
DB	4931	TCCCTCAGGATATAGTATGTTTCGTTTTCATAGGAGGGGAAATAGTCTTATGCA	4990
QY	3122	ATACTCTTGTAGTCTTGCAACATGTTAGTACGATGAGTACCAACATGCTTACAAGGAGAG	3181
DB	4991	ATACACTTGTAGTCTTGCAACATGTTAGTACGATGAGTACCAACATGCTTACAAGGAGAG	5050
QY	3182	AAAAGCACCGTGATGCGGATTTGGTGAAGTAAAGTACGATCGTGCCTTATTAGGA	3241
DB	5051	AAAAGCACCGTGATGCGGATTTGGTGAAGTAAAGTACGATCGTGCCTTATTAGGA	5110
QY	3242	AGGCAACAGACGGGTCTGACATGATGGAAGTAAAGTACGATCGTGCCTTATTAGGA	3300
DB	5111	AGGCAACAGACGGGTCTGACATGATGGAAGTAAAGTACGATCGTGCCTTATTAGGA	5170
QY	3301	ATTGTATTAAAGTCTAGTCTGATACAGCAACGCTTACCATTTACCAACATTTGGT	3360
DB	5171	ATTGTATTAAAGTCTAGTCTGATACAAATTAACCCCATTTGACCATTTACCAACATTTGGT	5230
QY	3361	GTGCACCTCCAGCTTACGCTGCGCAAGCACTCAGGCGCGCAAGGGCTGCTAAAGGAAG	3420
DB	5231	GTGCACCTCCAGCTTACGCTGCGCAAGCACTCAGGCGCGCAAGGGCTGCTAAAGGAAG	5290
QY	3421	CGGAACAGTGAAGACGAGTCCGAGAAACGGTCTGACCCCGGATGAATGTGAGTAC	3480
DB	5291	CGGAACAGTGAAGACGAGTCCGAGAAACGGTCTGACCCCGGATGAATGTGAGTAC	5350
QY	3481	TGGGCTATCTGGAACAGGAAAAACGCAAGCGCAAGAGAGGAGTAGCTTGCAATGGG	3540

DB	5351	TGGGCTATCTGGAACAGGAAAAACGCAAGCGCAAGAGAGGAGTAGCTTGCAATGGG	5410
QY	3541	CTTACATGGCGATAGCTAGACTGGCGGTTTATGAGCAGCAAGGACCGGAATTCGCA	3600
DB	5411	CTTACATGGCGATAGCTAGACTGGCGGTTTATGAGCAGCAAGGACCGGAATTCGCA	5470
QY	3601	GCTGGGGCGGCTCTGTTAAGTTGGGAAGCCCTGCAAAAGTAAACTGATGGCTTTCTTG	3660
DB	5471	GCTGGGGCGGCTCTGTTAAGTTGGGAAGCCCTGCAAAAGTAAACTGATGGCTTTCTTG	5530
QY	3661	CCGCCAAGGATCTGATGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGATCG	3720
DB	5531	CCGCCAAGGATCTGATGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGATCG	5590
QY	3721	TTTTCGATGATTGNAACAAGATGATTCACGCGAGGTTCTCCGCGCGCTTGGTGGAGAGG	3780
DB	5591	TTTTCGATGATTGNAACAAGATGATTCACGCGAGGTTCTCCGCGCGCTTGGTGGAGAGG	5650
QY	3781	CTATTTCGGCTATGATGCGGCACACAGACAATCGGCTGCTCTGATGCCCGCTGTTCCGG	3840
DB	5651	CTATTTCGGCTATGATGCGGCACACAGACAATCGGCTGCTCTGATGCCCGCTGTTCCGG	5710
QY	3841	CTGTACGCGCAGGCGGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAAT	3900
DB	5711	CTGTACGCGCAGGCGGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAAT	5770
QY	3901	GAACTGACGACGAGGCGCGGCTATCGTGGCTGGCCAGCAGCGGGGTTCTTTCGCA	3960
DB	5771	GAACTGACGACGAGGCGCGGCTATCGTGGCTGGCCAGCAGCGGGGTTCTTTCGCA	5830
QY	3961	GCTGTGCTCGACGTTGCTACTGAAGCGGGAAGGACTGGCTGCTATTGGGCGAAGTGCCG	4020
DB	5831	GCTGTGCTCGACGTTGCTACTGAAGCGGGAAGGACTGGCTGCTATTGGGCGAAGTGCCG	5890
QY	4021	GGGCAGGATCTCTGTGTCATCTCACCTTGTCTCTCCCGAGAAAGTATCCATCATGGCTGAT	4080
DB	5891	GGGCAGGATCTCTGTGTCATCTCACCTTGTCTCTCCCGAGAAAGTATCCATCATGGCTGAT	5950
QY	4081	GCAATGGCGGCTGTCATACGTTGATCCGGCTACCTGCGCATTCGACACCAAGCGGAA	4140
DB	5951	GCAATGGCGGCTGTCATACGTTGATCCGGCTACCTGCGCATTCGACACCAAGCGGAA	6010
QY	4141	CATCGCATCGAGGAGCAGCTACTCGGATGGAAGCGGCTCTTTCGATCAGGATGATCTG	4200
DB	6011	CATCGCATCG - GCAGCAGCTACTCGGATGGAAGCGGCTCTTTCGATCAGGATGATCTG	6069
QY	4201	GACGAAGAGCATCAGGGGCTCGCGCCACGCGAACTGTTCCGCAAGGCTCAAGGCGCGCATG	4260
DB	6070	GACGAAGAGCATCAGGGGCTCGCGCCACGCGAACTGTTCCGCAAGGCTCAAGGCGCGCATG	6129
QY	4261	CCCGACGGCAGGATCTGCTGCTGACCCCATGGCGATGCTGCTGCGCAATATCATGGTG	4320
DB	6130	CCCGACGGCAGGATCTGCTGCTGACCCCATGGCGATGCTGCTGCGCAATATCATGGTG	6189
QY	4321	GAAAATGGCGCTTTCTGGAATTCATCGACTGTGGCGGCTGGGTGGCGGACCGCTAT	4380
DB	6190	GAAAATGGCGCTTTCTGGAATTCATCGACTGTGGCGGCTGGGTGGCGGACCGCTAT	6249
QY	4381	CAGGACATAGCGTTGGCTAGCCGTCATTTGCTGAAGAGCTTGGCGGCGAATGGGCTGAC	4440
DB	6250	CAGGACATAGCGTTGGCTAGCCGTCATTTGCTGAAGAGCTTGGCGGCGAATGGGCTGAC	6309
QY	4441	CGCTTCTCGTCTTTACGGTATCGCCGCTCCCGAATTCGACGCGCATCGCTTCTATCGC	4500
DB	6310	CGCTTCTCGTCTTTACGGTATCGCCGCTCCCGAATTCGACGCGCATCGCTTCTATCGC	6369
QY	4501	CTTCTTACGAGTCTTCTGAGCGGGACTCTGGGGTTCGATA 4542	
DB	6370	CTTCTTACGAGTCTTCTGAGCGGGACTCTGGGGTTCGATA 6411	

RESULT 15

US-08-464-248A-6
; Sequence 6, Application US/08464248A
; Patent No. 5703209
; GENERAL INFORMATION:
; APPLICANT: Jacobeen, J. S.
; APPLICANT: Vittek, M. P.
; TITLE OF INVENTION: No. 5703209el Amyloid Precursor and Method of
; TITLE OF INVENTION: Using Same to Access Agents Which Down-Regulate Formation
; TITLE OF INVENTION: of B-Amyloid Peptide
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Cyanamid Company
; STREET: One Cyanamid Plaza
; CITY: Wayne
; STATE: New Jersey
; COUNTRY: United States
; ZIP: 07470-8426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,248A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Barnhard, Elizabeth M.
; REGISTRATION NUMBER: 31,088
; REFERENCE/DOCKET NUMBER: 31,844-02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201)831-3246
; TELEFAX: (201)831-3305
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8591 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2393..3868
US-08-464-248A-6

Query Match	28.0%;	Score 1448.4;	DB 2;	Length 8591;
Best Local Similarity	99.5%;	Pred. No. 0;		
Matches 1474;	Conservative 0;	Mismatches 6;	Indels 2;	Gaps 2;
Qy	3062	TCCCTCAGGATATAGTAGTTTCGCTTTTGCATAGGAGGGGNAATGTAGTCTTATGCA	3121	
Db	4931			
		TCCCTCAGGATATAGTAGTTTCGCTTTTGCATAGGAGGGGAAATGTAGTCTTATGCA	4990	
Qy	3122	ATACCTTTGTAGTCTTCCAAACATGGTAAACATAGATTAGCAACATGCCTTTACAAGAGAGAG	3181	
Db	4991	ATACCTTTGTAGTCTTCCAAACATGGTAAACATAGATTAGCAACATGCCTTTACAAGAGAG	5050	
Qy	3182	AAAAAGCACCGTGATCCGATTGGTGGGAAGTAAGGTGGTACGATCGTGCTTATTAGGA	3241	
Db	5051	AAAAAGCACCGTGATCCGATTGGTGGGAAGTAAGGTGGTACGATCGTGCTTATTAGGA	5110	
Qy	3242	AGGCAACAGACGGTCTTGACATGATTGGACGACCACTGAATTTCCGCATTGCAGAGAT-	3300	
Db	5111	AGGCAACAGACGGTCTTGACATGATTGGACGACCACTGAATTTCCGCATTGCAGAGATA	5170	
Qy	3301	ATTGTATTTAAGTGCCTAGCTCGATACAGCAAAACGCCATTGTGACCATTTACCACATTGGT	3360	
Db	5171	ATTGTATTTAAGTGCCTAGCTCGATACAAATAAACGCCATTGTGACCATTTACCACATTGGT	5230	
Qy	3361	GTGCACCTCCAGCTTCACGGTCCGCAAGCACTCAGGGGCGAAGGGCTGCTAAAGGAAG	3420	
Db	5231	GTGCACCTCTCTAGCTTCACGGTCCGCAAGCACTCAGGGGCGAAGGGCTGCTAAAGGAAG	5290	

Qy	3421	CGGAA	CACGTAGAAAGCCAGTCCGCGAGAAACCGTGTCTGACCCCGAGTGAATGTCAAGTAC	3448
Db	5291	CGGAA	CACGTAGAAAGCCAGTCCGCGAGAAACCGTGTCTGACCCCGAGTGAATGTCAAGTAC	5350
Qy	3481	TGGCGTATCTGGACAGGGAAACGACGCGCAAGAGAAAGCAGGTAGCTTGCAGTGGG	3540	
Db	5351	TGGCGTATCTGGACAGGGAAACGACGCGCAAGAGAAAGCAGGTAGCTTGCAGTGGG	5410	
Qy	3541	CTTACATGGCGATAGCTAGACTGGGCGGTTTTATATGGACAGCAAGCGAACTGGCAATTGCCA	3600	
Db	5411	CTTACATGGCGATAGCTAGACTGGGCGGTTTTATATGGACAGCAAGCGAACTGGCAATTGCCA	5470	
Qy	3601	GCTGGGCGGCCCTCTGTGTAAAGTTGGGAAGCCCTGCAAAGTAAACTGGATGGCTTCTTTG	3660	
Db	5471	GCTGGGCGGCCCTCTGTGTAAAGTTGGGAAGCCCTGCAAAGTAAACTGGATGGCTTCTTTG	5530	
Qy	3661	CCGCCAAGGATCTGATGGCGCACGGGATCAAGATCTGATCAAGACAGACAGATCAGAGATCG	3720	
Db	5531	CCGCCAAGGATCTGATGGCGCACGGGATCAAGATCTGATCAAGACAGACAGATCAGAGATCG	5590	
Qy	3721	TTTTCGATGATGAAACAAGATGATTCACGCGAGGTTCTCCGGCCGCTTGGGTGGAGAGG	3780	
Db	5591	TTTTCGATGATGAAACAAGATGATTCACGCGAGGTTCTCCGGCCGCTTGGGTGGAGAGG	5650	
Qy	3781	CTATTGGGCTATGACTGGGCAACAAGCAAACTCGGCTGCTCTGATGCGCGCGTGTCCCGG	3840	
Db	5651	CTATTGGGCTATGACTGGGCAACAAGCAAACTCGGCTGCTCTGATGCGCGCGTGTCCCGG	5710	
Qy	3841	CTGTACGCGCAGGGCGCGGTTCTTTTGTCAAGACCGACTGTCTCGGTGCGCTTCTTCGCGA	3900	
Db	5711	CTGTACGCGCAGGGCGCGGTTCTTTTGTCAAGACCGACTGTCTCGGTGCGCTTCTTCGCGA	5770	
Qy	3901	GAACCTGACGAGCAGGCGCGGCTATCGTGCTGCGCCACAGCAGGCGGTTCTTCGCGCA	3960	
Db	5771	GAACCTGACGAGCAGGCGCGGCTATCGTGCTGCGCCACAGCAGGCGGTTCTTCGCGCA	5830	
Qy	3961	GCTGTGCTCGACGTTGTCTACTGAAGCGGGAAGGGACTGGTGTCTATTGGGCGAAGTGC	4020	
Db	5831	GCTGTGCTCGACGTTGTCTACTGAAGCGGGAAGGGACTGGTGTCTATTGGGCGAAGTGC	5890	
Qy	4021	GGGCGAGATCTCTGTCACTCACTTGTCTCTGCGGAGAAAGTATCCATCATGGCTGAT	4080	
Db	5891	GGGCGAGATCTCTGTCACTCACTTGTCTCTGCGGAGAAAGTATCCATCATGGCTGAT	5950	
Qy	4081	GCAATGGCGGCTGCATAGCTTGATTCGGCTACTCTGCCCATTCGACACCAACGCGAAA	4140	
Db	5951	GCAATGGCGGCTGCATAGCTTGATTCGGCTACTCTGCCCATTCGACACCAACGCGAAA	6010	
Qy	4141	CATCGCATCGACGACGATCTCGATGGAAGACCGGCTTGTCTGATCAGGATGATCTG	4200	
Db	6011	CATCGCATCG - GCGAGCAGTACTCGGATGGAAGCCGGTCTTGTCTGATCAGGATGATCTG	6069	
Qy	4201	GACGAGAGCATCAGGGGCTCGCGCAGCGGACTGTTGCGCAGGCTCAAGGCGCGCATG	4260	
Db	6070	GACGAGAGCATCAGGGGCTCGCGCAGCGGACTGTTGCGCAGGCTCAAGGCGCGCATG	6129	
Qy	4261	CCCGACGGCAGGATCTCGTGTGACCCATGGCGATGCTGCTTGGCGAATATCATGGTG	4320	
Db	6130	CCCGACGGCAGGATCTCGTGTGACCCATGGCGATGCTGCTTGGCGAATATCATGGTG	6189	
Qy	4321	GAAAAATGGCGCTTTTCTGGAATTCATCGACTGTGGCGGCTGGGTTGGCGGACCGCTAT	4380	
Db	6190	GAAAAATGGCGCTTTTCTGGAATTCATCGACTGTGGCGGCTGGGTTGGCGGACCGCTAT	6249	
Qy	4381	CAGGACATAGCTTGCTACCGGTGATATTGCTGAAGAGCTTGGCGGCAATGGGCTGAC	4440	
Db	6250	CAGGACATAGCTTGCTACCGGTGATATTGCTGAAGAGCTTGGCGGCAATGGGCTGAC	6309	
Qy	4441	CGCTTCTCTGTGCTTACCGATTCGCGGCTTCCCGATTCGACGCGCATTCGCTTCTATCGC	4500	
Db	6310	CGCTTCTCTGTGCTTACCGATTCGCGGCTTCCCGATTCGACGCGCATTCGCTTCTATCGC	6369	
Qy	4501	CTTCTTTGACAGGTTCTTCTGACGGGACTCTGGGGTTGATA	4542	

Db 6370 CTTCTTGACGAGTTCTTCTGAGCGGACTCTGGGGTTCGAAA 6411
|||||

Search completed: February 27, 2006, 08:49:42
Job time : 893 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 27, 2006, 08:49:57 ; Search time 6646 Seconds
(without alignments)
3421.213 Million cell updates/sec

Title: US-10-618-570-2

Perfect score: 2059

Sequence: 1 MAQWDDPPDQEDTDSCTES.....DWKATRVGINIFRLRTQKE 400

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWEB.apool/US10618570/runat 27022006 063854 13050/app query.fasta_1
-DB=GenEmbl -OPMT=fastrap -SUPFIX=rge -MINMATCH=0.1 -LOFPCU=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02h
-USER=US10618570 @CGN 1 1 4939 @runat 27022006 063854 13050 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WAE TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_px.*

9: gb_pro.*

10: gb_ats.*

11: gb_sy.*

12: gb_un.*

13: gb_vi.*

14: gb_btg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2059	100.0	5177	6	BD136828 Biotin-bi
2	2059	100.0	5177	6	AX018983 Sequence
3	1381.5	67.1	1588	6	AR365560 Sequence

4	1381.5	67.1	1807	6	AR365559	AR365559 Sequence
5	1381	67.1	1592	4	BTSICAVR1	X51689 B.taurus mr
6	1381	67.1	2135	4	BTSICAVR2	X54183 B.taurus mr
7	1042.5	50.6	1167	8	AF037351	AF037351 Homo sapi
8	1042.5	50.6	1347	6	E03268	E03268 cDNA sequen
9	1042.5	50.6	1347	6	E05211	E05211 DNA encodin
10	1042.5	50.6	1347	6	BD000149	BD000149 Anthuman
11	1042.5	50.6	1367	6	CQ722728	CQ722728 Sequence
12	1042.5	50.6	1367	6	I72357	I72357 Sequence 3
13	1042.5	50.6	1367	6	AR270697	AR270697 Sequence
14	1042.5	50.6	1367	8	HUMPHSR2	D90188 Homo sapien
15	1042.5	50.6	2028	6	E03267	E03267 cDNA sequen
16	1042.5	50.6	2028	6	E05210	E05210 DNA encodin
17	1042.5	50.6	2028	6	I72356	I72356 Sequence 1
18	1042.5	50.6	2028	6	BD000148	BD000148 Anthuman
19	1042.5	50.6	2028	8	HUMPHSR1	D90187 Homo sapien
20	1042.5	50.6	2209	8	BC063878	BC063878 Homo sapi
21	1042.5	50.6	3704	6	AX281714	AX281714 Sequence
22	1039.5	50.5	2037	6	I41352	I41352 Sequence 1
23	994.5	48.3	1330	6	AX926628	AX926628 Sequence
24	994.5	48.3	1330	6	AX926677	AX926677 Sequence
25	994.5	48.3	1365	4	RABMSRT1	D13381 Oryctolagus
26	987.5	48.0	1563	4	RABMSRT1	L11692 Oryctolagus
27	987.5	48.0	1862	4	RABMSRT1	L11693 Oryctolagus
28	819	39.8	1425	9	AF203781	AF203781 Mus muscu
29	801	38.9	1065	9	MUSMSRT2	D13382 Mus sp. msn
30	801	38.9	1430	9	MUSMSRT1	L04274 Mus musculu
31	801	38.9	1572	9	BC003814	BC003814 Mus muscu
32	801	38.9	1946	9	MUSMSRT1	L04275 Mus musculu
33	710.5	34.5	257208	14	AC157118	AC157118 Bos tauru
34	688	33.4	459	6	CQ830753	CQ830753 Sequence
35	678.5	33.0	1746	6	CS104130	CS104130 Sequence
36	678.5	33.0	1746	11	AJ966780	AJ966780 Synthetic
37	678	32.9	604	5	GGAVIR	X05343 Chicken mRN
38	678	32.9	604	6	A93651	A93651 Sequence 3
39	678	32.9	604	6	AR177253	AR177253 Sequence
40	678	32.9	604	6	AR217937	AR217937 Sequence
41	677.5	32.9	873	6	CS104134	CS104134 Sequence
42	586	28.5	897	6	CS104131	CS104131 Sequence
43	586	28.5	897	11	AJ616762	AJ616762 Synthetic
44	537	26.1	1119	10	CHKAVID	L27818 Gallus gall
45	523	25.4	1224	5	GGA311647	AJ311647 Gallus ga

ALIGNMENTS

RESULT 1	LOCUS	BD136828	BD136828	Biotin-binding receptor molecule.	DNA	linear	PAT 18-SEP-2002
BD136828	DEFINITION	BD136828	BD136828	Biotin-binding receptor molecule.	5177 bp		
	ACCESSION	BD136828	BD136828	Biotin-binding receptor molecule			
	VERSION	BD136828.1	GI:23231773	Patent: JP 2002504328-A 1 12-FEB-2002;			
	KEYWORDS	JP 2002504328-A/1.		ARC THERAPEUTICS LTD			
	SOURCE	unidentified		OS Unidentified			
	ORGANISM	unclassified.		PN JP 2002504328-A/1			
	REFERENCE	1 (bases 1 to 5177)		PD 12-FEB-2002			
	AUTHORS	Herttua, S.V., Kulomaa, M., Lehtolainen, P., Marjomaki, V. and Airenne, K.		PF 23-FEB-1999 JP 2000532517			
	TITLE	Biotin-binding receptor molecule		PR 23-FEB-1998 GB 9803757.5, 24-JUN-1998 GB 9813653.4 PI			
	JOURNAL	ARC THERAPEUTICS LTD		SEPPO YLA HERTTUALA, MARKKU KULOMAA, PAULIINA LEHTOLAINEN, VAREPU PI			
	COMMENT	OS Unidentified		MARJOMAKI, KARI, AJRENNNE			
		PN JP 2002504328-A/1		PI C12N15/09, A61K48/00, A61P43/00, C07K14/705//A61K38/00, C12N15/00,			
		PD 12-FEB-2002		PC C12N15/09, A61K48/00, A61P43/00, C07K14/705//A61K38/00, C12N15/00,			
		PF 23-FEB-1999 JP 2000532517		PC A61K37/02			
		PR 23-FEB-1998 GB 9803757.5, 24-JUN-1998 GB 9813653.4 PI		CC Strandedness: Single;			


```

Best Local Similarity: 100.0%      Mismatches: 0
Query Match: 100.0%      Indels: 0
DB: 6      Gaps: 0

US-10-618-570-2 (1-400) x AX018983 (1-5177)

Qy 1 MetAlaGlnTTPAspAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
Db 1071 ATGGCAGAGTGGGATGACCTTCCTCGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCT 1130
Qy 21 VallysPheAspAlaAArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 1131 GTGAAGTTCGATGCTCGCTCAGTGACAGCTTGGCTTCCTCCCATCTCTAAAAATGGGCCA 1190
Qy 41 ThrLeuGlnGluAArgMetLysSerTyrLysThrAlaLeuLeuLeuLeuLeuVal 60
Db 1191 ACTCTTCAAGAGAGGATGAAGCTTATAAACTGCACCTGATCACCTTTATCTCATTTGTG 1250
Qy 61 PheValValLeuValProIleGlyLeValAlaAlaGlnLeuLeuLeuLeuLeuThr 80
Db 1251 TTTGTAGTTCCTGCGCCCATCATTTGGCATATGCGAGCTCAGCTCCTGAAATGGGAACG 1310
Qy 81 LysAsnCysThrValGlySerValAsnAlaAspIleSerProSerProGluGlyLysGly 100
Db 1311 AAGAATTGCACGGTGGCTCAGTTAATGCAGATATATCTCCAGTCCGGAAGGCAAGGA 1370
Qy 101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluAArgMetSerAsnMet 120
Db 1371 AATGCGAGTGAAGATGAATGAGATTTTCGAAAGCTGTGATGGAACGATCAGCAACATG 1430
Qy 121 GluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuLeuLeuLeuPhe 140
Db 1431 GAAAGCAGAAATCCAGTATCTTTTCAGATAATGAAGCCAAATCTCTAGATGCTTAAGAATTC 1490
Qy 141 GlnAsnPheSerIleThrThrAspGlnAArgPheAsnAspValLeuPheGlnLeuAsnSer 160
Db 1491 CAAATTTTCAGCAATACACTGATCAAGATTTAATGATGTTCTTTCCAGCTAAATTC 1550
Qy 161 LeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeuVal 180
Db 1551 TTACTTTCTCCATCCAGGAACATGAGATATATATAGGGATATCTCCAAGTCATTAGTA 1610
Qy 181 GlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal 200
Db 1611 GGTCGTAACACACAGTACTTGAATTTGAGTTTCACTATTTGAACACTCAATGGCAGAGTC 1670
Qy 201 GlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluAArgIleTyrAsn 220
Db 1671 CAAGAGATGCATTTAAACAACAAGAGAGATGCGTAAATTAGAGGAGCGTATATACAAT 1730
Qy 221 AlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIleLys 240
Db 1731 GCATCAGCAGAAATTAAGTCTCTAGATGAATAAACAAGTATATTTGGACAGGAATAAATA 1790
Qy 241 GlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTyrGluHis 260
Db 1791 GGGGAAATGAAACTGTTGAATAATATCACTAATGATCTGAGGCTGAAGGATGGGAACAT 1850
Qy 261 SerGlnThrLeuLysAsnIleThrLeuLeuGlnGlnAlaAArgLysCysSerLeuThrGly 280
Db 1851 TCTCAGACATTTGAATAATATCACTTTACTCAAGGTGCCAAGATGCTCGCTCACTGGG 1910
Qy 281 LysTyrThrAsnAspLeuGlySerAsnMetThrIleGlyAlaValAsnSerArgGlyGlu 300
Db 1911 AAATGGACCAACGATCTGGGCTCCACATGACCATCGGGGTGTGNAACAGAGGTGAA 1970
Qy 301 PheThrGlyThrTyrIleThrAlaValThrAlaThrSerAsnGluIleLysGluSerPro 320
Db 1971 TTCAAGGCACCTACATCAGACCGGTAAACGCCACATCAATAGATCAAGAGAGTCACCA 2030
Qy 321 LeuHisGlyThrGlnAsnThrIleAsnLysArgThrGlnProThrPheGlyPheThrVal 340
Db 2031 CTGCATGGGACACAAAACACCATCAACAAGAGGCCACCGCCACCTTTGGCTTCACCGTC 2090

```

```

Qy 341 AsnTyrLysPheSerGluSerThrThrValPheThrGlyGlnCysPheIleAspArgAsn 360
Db 2091 AATTGGAAGTTTTCAGAGTCCACCAGTCTCTTCAAGGGCCAGTCTTTCATAGACAGGAAT 2150
Qy 361 GlyLysGluValLeuLeuLysThrMetTyrLeuLeuArgSerValAsnAspIleGlyAsp 380
Db 2151 GGGGAAGAGGTCCTGAAGACCATGTGCTGTCGGGTCAAGTGTTAATGACATTTGGTGT 2210
Qy 381 AspTyrLysAlaThrArgValGlyIleAsnIlePheThrArgLeuAArgThrGlnLysGlu 400
Db 2211 GACTGGAAAGCTACAGGTCGGCATCAACATCTTCACTCGCTCGGCACACAGAGGAG 2270

RESULT 3
AR365560 1588 bp DNA linear PAT 03-SEP-2003
LOCUS Sequence 3 from patent US 5510466.
DEFINITION AR365560
ACCESSION AR365560.1 GI:34429331
VERSION AR365560.1
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1588)
AUTHORS Krieger, M. and Kodama, T.
TITLE Scavenger receptor protein and antibody thereto
JOURNAL Patent: US 5510466-A 3 23-APR-1996;
Massachusetts Institute of Technology; Cambridge, MA
FEATURES
    source          1..1588
                    /organism="unknown"
                    /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 3,31e-89      Length: 1588
Score: 1381.50      Matches: 307
Percent Similarity: 79.4%      Conservative: 24
Best Local Similarity: 73.6%      Mismatches: 66
Query Match: 67.1%      Indels: 22
DB: 6      Gaps: 6

US-10-618-570-2 (1-400) x AR365560 (1-1588)

Qy 1 MetAlaGlnTTPAspAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
Db 4 ATGGCAGAGTGGGATGACCTTCCTCGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCT 63
Qy 21 VallysPheAspAlaAArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 64 GTGAAGTTCGATGCTCGCTCAGTGACAGCTTTGCTTCTCCCATCTCTAAAAATGGGCCA 123
Qy 41 ThrLeuGlnGluAArgMetLysSerTyrLysThrAlaLeuLeuLeuLeuLeuVal 60
Db 124 ACTCTTCAAGAGAGGATGAAGTCTTATAAACTGCACCTGATCACCTTTATCTCATTTGTG 183
Qy 61 PheValValLeuValProIleGlyLeValAlaAlaGlnLeuLeuLysTyrGluThr 80
Db 184 TTTGTAGTTCCTGCGCCCATCATTTGGCATATGTCAGTATGTCAGCTCAGCTCTGAAATGGGAACG 243
Qy 81 LysAsnCysThrValGlySerValAsnAlaAspIleSerProSerProGluGlyLysGly 100
Db 244 AAGAATTTGCACGGTGGCTCAGTTAATGCAGATATATCTCCAGTCCGGAAGGCAAGGA 303
Qy 101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluAArgMetSerAsnMet 120
Db 304 AATGGCAGTGAAGATGAATGAGATTTTCGAAAGCTGTGATGGAACGATGAGCAACATG 363
Qy 121 GluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPhe 140
Db 364 GAAAGCAGATTCAGTATCTTTTCAGATATGAGGCCAATCTCTAGATGCTTAAGAATTC 423
Qy 141 GlnAsnPheSerIleThrThrAspGlnAArgPheAsnAspValLeuPheGlnLeuAsnSer 160

```

```
Db 424 CAAAATTTCCAGTAAACAACATGATCAAGATTTAATGATGTTCTTTCCAGTAAATTC 483
Qy 161 LeuLeuSerSerIleGlnGluHieGluAsnIleIleGlyAspIleSerLysLeuVal 180
Db 484 TTACTTTCTCTCCATCCAGGAACATGAGAAATATCATAGGGGATATCTCAAGTCAATAGTA 543
Qy 181 GlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal 200
Db 544 GGTCTGAACACCAACAGTACTTGTATTTGCAGTTTCAGATTGAAACACTGAATGGCAGAGTC 603
Qy 201 GlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluAArgIleTyrAsn 220
Db 604 CAAGAGAATGCATTTAAACAACAAGAGAGATGCGTAAATTAGAGGCGGTATATACAAT 663
Qy 221 AlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIleLys 240
Db 664 GCATCAGCAGAAATTAAGTCTCTAGATGAAACCAAGTATATATTTGGAAACAGGAATAAAA 723
Qy 241 GlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTyrGluHis 260
Db 724 GGGGAATGAACCTGTGTAATATATCTAATGATCTGAGGCTGAAGGATTGGGAACAT 783
Qy 261 SerGlnThrLeuLysAsnIleThrLeuLeuGlnGlyAla----- 273
Db 784 TCTCAGACATTTGAAAAATATACATTTACTCCAGGTCCTCTCGACCTCCAGGTTGAAAAA 843
Qy 274 -----ArgLysCysSerLeuThrGlyLysThrAsnAspLeuGly---SerAsnMet 290
Db 844 GGAGATAGAGGCCCTCTCTGGACA---AAATGGTATACAGGCTTTCCAGGCTCAATAGG 899
Qy 291 ThrIleGlyAlaValAsnSerArgGlyGluPheThrGlyThrTyrIleThrAlaValThr 310
Db 900 TACTCCAGGCTTTAAAGTGATCGGGGATCTCTGGTTTACTCGAGTTCAGGATTCCTCC 959
Qy 311 AlaThrSerAsnGluIleLysGluSer-ProLeuHieGlyThrGlnAsnThrIleAsnLy 330
Db 960 AGGACCAATGGGAAGACCGGGAAGCCAGGACTTAATGACAAAAAGGCCAGAGGGGAGA 1019
Qy 330 sArgThrGlnProThrPhe-----GlyPheThrVal-AsnTyrLysPheSerGlu 347
Db 1020 AAAACCCACTCCAAACCACTCCAAACACATCTAATACAGTCCGACTGGTGGTGCGCGG 1079
Qy 347 erThrThrValPheThrGlyGlnCysPheIleAspArgAsnGlyLysGluValLeuLys 367
Db 1080 CCCTCAGCAAGGCAGAGTGGAGATTTTTCAGAAAGGCAGTGGGGTACGGTGTGTGACGA 1139
Qy 367 hrMetTyrLeuLeuArgSerSerVal-AsnAspIleGlyAsp---AspTyrLysAlaThr 385
Db 1140 CCGC-TGGGAACCTGGTGGAGGACTGGTTCGTCGACAGGAGCTTGGGATACAAAGTGTTC 1198
Qy 386 ArgValGlyIleAsn-----IlePheThrArgLeuArgThrGln 398
Db 1199 AAAGTGTGCATPAGCGAGCTTATTTTGGAAAAAGGTACGGGTCCAA 1243

RESULT 4
AR365559 1807 bp DNA linear PAT 03-SBP-2003
LOCUS
DEFINITION Sequence 1 from patent US 5510466.
ACCESSION AR365559
VERSION AR365559.1 GI:34429330
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
1 (bases 1 to 1807)
REFERENCE Krieger, M. and Kodama, T.
AUTHORS Scavenger receptor protein and antibody thereto
TITLE Patent: US 5510466-A 1 23-APR-1996;
JOURNAL Massachusetts Institute of Technology; Cambridge, MA
FEATURES
Location/Qualifiers
1..1807
/organism="unknown"
```

```
ORIGIN
/mol_type="genomic DNA"
Alignment Scores: 3.79e-89 Length: 1807
Pred. No.: 1391.50 Matches: 288
Score: 85.2% Conservative: 16
Percent Similarity: 80.7% Mismatches: 34
Best Local Similarity: 67.1% Indels: 20
Query Match: 6 Gaps: 4
DB: 6

US-10-618-570-2 (1-400) x AR365559 (1-1807)
Qy 1 MetAlaGlnTyrAspAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 36 ATGGCACAGTGGGATGACTTTCCTGATCAGCAAGAGAGACACTGCACAGTGTACAGAGTCT 95
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProProHisProLysAsnGlyPro 40
Db 96 GTGAAGTTCGATGCTCGCTCAGTGACAGCTTTGCTTCTCCCATCTCTAAAAATGGCCCA 155
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60
Db 156 ACTCTTCAAGAGAGGATGAAGTCTTATAAAACAGCTGATCACCCTTTATCTCATTTGTG 215
Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTyrGluThr 80
Db 216 TTTGTAGTCTCTGCGCCATCATTTGGCATAGTGGCAGCTCAGCTCTCTGAAATGGGAAACG 275
Qy 81 LysAsnCysThrValGlySerValAsnAlaAspIleSerProSerProGluGlyLysGly 100
Db 276 AAGAATTCACCGTGGCTCAGTTAATGATGATATATCTCAAGTCCGGAAGGCAAAAGGA 335
Qy 101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120
Db 336 AATGGCAGTGAAGATGAATGAGATTTTCGAGAAGCTGTGATGGAACCATGAGCAACATG 395
Qy 121 GluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPhe 140
Db 396 GAAAGCAGATCCAGTATCTTTCAGATATGAGGCCAATCTCTAGATGCTNAGAAATTC 455
Qy 141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
Db 456 CAAAATTTCCAGCAATAACAAGTGAATTAATGATGATGATGATGATGATGATGATGATGAT 515
Qy 161 LeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeuVal 180
Db 516 TTACTTTCTCTCCATCCAGGAACATGAGAAATATCATAGGGGATATCTCCAAGTCAATAGTA 575
Qy 181 GlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal 200
Db 576 GGTCTGAACACCAACAGTACTTGTATTTGCACTTCAGTATTCAGTATTCAGTATTCAGT 635
Qy 201 GlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluAArgIleTyrAsn 220
Db 636 CAAGAGAATGCATTTAAACAACAAGAGAGATGCGTAAATTAGAGGAGCGGTATATACAAT 695
Qy 221 AlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIleLys 240
Db 696 GCATCAGCAGAAATTAAGTCTCTAGATGAAACCAAGTATATATTTGGAAACAGGAATAAAA 755
Qy 241 GlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTyrGluHis 260
Db 756 GGGGAATGAACCTGTGTAATATATCTAATGATCTGAGGCTGAAGGATTGGGAACAT 815
Qy 261 SerGlnThrLeuLysAsnIleThrLeuLeuGlnGlyAla----- 273
Db 816 TCTCAGACATTTGAAAAATATCACTTTACTCCAGGTCCTCTCGACCTCCAGGTTGAAAAA 875
Qy 274 -----ArgLysCysSerLeuThrGlyLysThrAsnAspLeuGly---SerAsnMet 290
Db 876 GGAGATAGAGGCCCTCTCTGGACA---AAATGGTATACAGGCTTTCCAGGCTTAATAGG 931
```

```

Qy 291 ThrileGlyAlaValAsnSerArgGlyCluPhethrGlyThrIleThrAlaValThr 310
Db 932 TACTCCAGGTCTTAAAGGTGATCGGGGATCTCGTTTACCTGGAGTTCGAGGATCCC 991
Qy 311 AlaThrSerAsnGluLeuGluSerProLeuHisGlyThrGlnAsnThrIleAsnLys 330
Db 992 AGGACCATGGGAAGACCGGAGCCAGGACTATTGGACAAAAGCCAGAGGGGAGA 1051
Qy 331 ArgThrGlnProThrPheGlyPheThrValAsnTrpLysPheSerGluSer 347
Db 1052 AAAAGGGAG-----TGGAGCATGCAAGACC 1078

RESULT 5
BTSCAVR1
LOCUS B.taurus mRNA 1592 bp mRNA linear MAM 18-APR-2005
DEFINITION B.taurus mRNA for macrophage scavenger receptor type I.
ACCESSION X51689 X54182
VERSION X51689.1 GI:734
KEYWORDS lipid metabolism; macrophage scavenger receptor type I; membrane glycoprotein.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
REFERENCE 1 (bases 1 to 1592)
AUTHORS Kodama,T.; Freeman,M., Rohrer,L., Zabrecky,J., Matsudaira,P. and Krieger,M.
TITLE Type I macrophage scavenger receptor contains alpha-helical and collagen-like coiled coils
JOURNAL Nature 343 (6258), 531-535 (1990)
PUBMED 2300204
REFERENCE 2 (bases 1 to 1592)
AUTHORS Krieger,M.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-1990) Krieger M., Massachusetts Institute of Technology, Room E25-236, 77 Massachusetts Ave, Cambridge MA 02139, U S A

FEATURES
source
1..1592
location/Qualifiers
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="caxon:9913"
/clone="pBSR7"
/tissue type="lung"
/clone lib="1ZAPII"
22..1383
/codon_start=1
/product="scavenger receptor type I"
/protein_id="CAA35987.1"
/db_xref="GI:735"
/db_xref="GOA:P21758"
/db_xref="InterPro:IPR001190"
/db_xref="InterPro:IPR003543"
/db_xref="InterPro:IPR008160"
/db_xref="InterPro:IPR008161"
/db_xref="UniProt/Swiss-Prot:P21758"
/translaton="MAOWDDPDDQEDTSCVESKFDARSVTALLPPHPKNGPTLQES
RKSKYKTAIIYLIVFVLPVPIIGIVAQLLKWKNTCTGVSYNADISPSPEKGNG
SEDMRFRFEANWMSRISQYLSNEANLLDAKQNFNSITTDQRFNDVLFQNS
LLSQIHEHNIIGDISVLGNTVLDLQFSIETINGRQENAFQOQENRKLBERI
YNASAKIKLEQVLEQEIKGEMKLNNITNDLRKDWHSQTLKNTITLLQPPGP
PGEKGRDPPQNGIPGPGIIGTGLKGDGRIISGLPGVPGFPGMGKTKPGKLGOK
GQKEKSGSMQRSNTVRLVGGSGPHEGRVEIFHEGOWGTVCDDRWELRGGLVVCES
LYKGVQSVHRAFGKGTGTPILNWFVFCFKESSIECRIRQGVGRACSHDEDAVGT
CTT"

CDS
3..68-89 Length: 1592
Pred. No.: 1384.00 Matches: 272
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0

ORIGIN
Alignment Scores:
Pred. No.: 1592
Score: 272
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Mismatches: 0

```

```

Query Match: 67.1% Indels: 0
DB: 4 Gaps: 0
US-10-618-570-2 (1-400) x BTSCAVR1 (1-1592)
Qy 1 MetAlaGlnTrpAspAspPheProAspGlnGlnuAspThrAspSerCysThrGluSer 20
Db 22 ATGGCAGAGTGGGATGACTTTCTGATCAGCAGAGGACACTGACAGCTGTACAGAGTCT 81
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 82 GTGAAGTTTCGATGCTCGCTCAGTCAGACAGCTTTGCTCTCCCATCTCTAAATGGCCCA 141
Qy 41 ThrLeuGlnGluArgMetLysSerTyrlsThrAlaLeuLeuThrLeuTyrlsLeuVal 60
Db 142 ACTCTTCAAGAGAGAGTGAAGTCTTATAAACTGCACCTGATCACCCTTTATCTCATGTG 201
Qy 61 PheValValLeuValProIleGlyIleValAlaAlaGlnLeuLeuTyrlsGluThr 80
Db 202 TTTGTAGTTCTCGTGGCCCATCATTTGGCATAGTGGCAGCTCAGCTCTCTCAATGGGAAACG 261
Qy 81 LysAsnCysThrValGlySerValAsnAlaAspIleSerProSerProGluGlyLysGly 100
Db 262 AAGAAATTCACGGTGGCTCAGTTAATGCAGATATATCTCCAGTCCGGAAGGCAAGGA 321
Qy 101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120
Db 322 AATGCCAGTGAAGATGAATGAGATTTTCGAGAAGCTGTGATGGAACGCATGAGCAACATG 381
Qy 121 GluSerArgIleGlnTyrlsLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPhe 140
Db 382 GAAAGCAGAAATCCAGTATCTTTTCAGATAATGAAGCCCAATCTCTTAGATGCTTAAGAAATTTTC 441
Qy 141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
Db 442 CAAAATTTTCAGCATACACATGATCAAGATTTTAAATGATGTCTTTTCCAGCTAAATTC 501
Qy 161 LeuLeuSerSerIleGlnGluHisGluAsnIleGlyAspIleSerLysSerLeuVal 180
Db 502 TTACTTTCTCCATCCAGGAACATGAGATATATATAGGGGATATCTCCAGCTATTAGTA 561
Qy 181 GlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal 200
Db 562 GGTCTGAACACACAGACTTGTATTTGAGTTCAGTATTGAAACACTGAATGGCAGAGTCT 621
Qy 201 GlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyrlsAsn 220
Db 622 CAAGAGATGCAATTTAAACAACAGAGAGATGCTTAATTTAGAGGAGCGTATATACAT 681
Qy 221 AlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrlsLeuGlnGlnLeuLys 240
Db 682 GCATCAGCAGAAATTAAGTCTCTAGATGMAAAACAAGTATATTTTGGAAACAGGMAATAAAA 741
Qy 241 GlyLysMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGluHis 260
Db 742 GGGGAAATGAACCTGTTGAATAATATCACTAATGATCTGAGGCTCAAGGATTTGGGAACAT 801
Qy 261 SerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 802 TCTCAGACATTTGAAAAATATACCTTTTACTCCAGGT 837

RESULT 6
BTSCAVR2
LOCUS B.taurus mRNA 2135 bp mRNA linear MAM 18-APR-2005
DEFINITION B.taurus mRNA for macrophage scavenger receptor, type II.
ACCESSION X54183
VERSION X54183.1 GI:736
KEYWORDS lipid metabolism; macrophage scavenger receptor type I; membrane glycoprotein.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
REFERENCE 1 (bases 1 to 2135)
AUTHORS Kodama,T.; Freeman,M., Rohrer,L., Zabrecky,J., Matsudaira,P. and Krieger,M.
TITLE Type II macrophage scavenger receptor contains alpha-helical and collagen-like coiled coils
JOURNAL Nature 343 (6258), 531-535 (1990)
PUBMED 2300204
REFERENCE 2 (bases 1 to 2135)
AUTHORS Krieger,M.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-1990) Krieger M., Massachusetts Institute of Technology, Room E25-236, 77 Massachusetts Ave, Cambridge MA 02139, U S A

FEATURES
source
1..2135
location/Qualifiers
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="caxon:9913"
/clone="pBSR7"
/tissue type="lung"
/clone lib="1ZAPII"
22..1383
/codon_start=1
/product="scavenger receptor type I"
/protein_id="CAA35987.1"
/db_xref="GI:735"
/db_xref="GOA:P21758"
/db_xref="InterPro:IPR001190"
/db_xref="InterPro:IPR003543"
/db_xref="InterPro:IPR008160"
/db_xref="InterPro:IPR008161"
/db_xref="UniProt/Swiss-Prot:P21758"
/translaton="MAOWDDPDDQEDTSCVESKFDARSVTALLPPHPKNGPTLQES
RKSKYKTAIIYLIVFVLPVPIIGIVAQLLKWKNTCTGVSYNADISPSPEKGNG
SEDMRFRFEANWMSRISQYLSNEANLLDAKQNFNSITTDQRFNDVLFQNS
LLSQIHEHNIIGDISVLGNTVLDLQFSIETINGRQENAFQOQENRKLBERI
YNASAKIKLEQVLEQEIKGEMKLNNITNDLRKDWHSQTLKNTITLLQPPGP
PGEKGRDPPQNGIPGPGIIGTGLKGDGRIISGLPGVPGFPGMGKTKPGKLGOK
GQKEKSGSMQRSNTVRLVGGSGPHEGRVEIFHEGOWGTVCDDRWELRGGLVVCES
LYKGVQSVHRAFGKGTGTPILNWFVFCFKESSIECRIRQGVGRACSHDEDAVGT
CTT"

```

[illegible]


```

Db      287 AAGAATTGCTCAGTTAGTTCACTAAATGCAAAATGATATAACTCAAGAGTCTCACGGGAAA 346
Qy      100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db      347 GGAATGACACGCGAAGGAAGAAATGAGATTTCAAGAAGTCTTTATGGAAACACATGAGCAAC 406
Qy      120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLeuAsn 139
Db      407 ATGAGAAGAGAAATCCAGCATATTTAGACATGGAAGCCCACTCATGGACACAGAGCAT 466
Qy      140 PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db      467 TTCAAAATTTTCAGCATGACCACTGATCAAAAGATTTAATGACATTTCTTCGACGTAAGT 526
Qy      160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleGlyAspIleSerLysSerLeu 179
Db      527 ACCTTGTCTTCTCAGTCCAGGGACATGGGAATGCAATAGATGAAATCTCCAAAGTCTCTTA 586
Qy      180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db      587 ATAAAGTTTGAATACCACTTGTGATTTTCAGCTCAACATAGAAATCTGAAATGGCAAA 646
Qy      200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db      647 ATCCAGAGAAATACCTTCAAAACACACAGAGAAATCAGTAATATTAGAGGCGGTGTAC 706
Qy      220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
Db      707 AATGTATCAGCAGAAATTTATGGCTATGAAAGAGAACAAAGTGCATTTTGGAAACAGGAAATA 766
Qy      240 LysGlyGluMetLysLeuAsnIleThrAsnAspLeuArgLysLeuGluGluArgIleTyr 259
Db      767 AAAGAGAGTGAAGTGAATGAAATGCAATAGATGAAATCTCCAAAGTCTCTTA 826

Qy      260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db      827 CATTCACAGACCTTGAGAAATATACCTTTAATTCAGGT 865

RESULT 9
E05211 LOCUS E05211 1347 bp RNA linear PAT 29-SEP-1997
DEFINITION DNA encoding human scavenger receptor II.
ACCESSION E05211
VERSION E05211.1 GI:2173401
KEYWORDS JP 1993192179-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE 1 (bases 1 to 1347)
AUTHORS Kodama,T., Matsumoto,A. and Suzuki,H.
TITLE ANTI-HUMAN SCAVENGER RECEPTOR ANTIBODY
JOURNAL Patent: JP 1993192179-A 2 03-AUG-1993;
          CHUGAI PHARMACEUT CO LTD
COMMENT OS Homo sapiens (human)
        PN JP 1993192179-A/2
        PD 03-AUG-1993
        PF 16-AUG-1991 JP 1991229728
        PR 27-AUG-1990 JP 90P 222398
        PI KODAMA TATSUHIKO, MATSUMOTO AKIYO, SUZUKI HIROSHI PC
        C12P21/08.G01N33/53.G01N33/577//A61K39/395.A61K39/395.C12N5/20, PC
        C12N15/06,
        PC (C12P21/08,C12N1:91);
        CC strandedness: Double;
        CC topology: Linear;
        FH Key Location/Qualifiers
        FT CDS 47..1123
        FT /product='human scavenger receptor II'.
        FT Location/Qualifiers
        FT 1..1347
        FT /organism='Homo sapiens'
FEATURES
source

```

```

/mol_type="genomic RNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 3.64e-65 Length: 1347
Pred. No.: 1042.50 Matches: 197
Score: 87.2% Conservative: 41
Percent Similarity: 72.2% Mismatches: 34
Best Local Similarity: 72.2% Indels: 1
Query Match: 6 Gaps: 1
DB: 6

US-10-618-570-2 (1-400) x E05211 (1-1347)

Qy      1 MetAlaGlnTrpAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
Db      47 ATGGAGCAGTGGGATCACTTTTCAAAATCAACAGAGGACACTGATAGCTCTCGAATCT 106
Qy      21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db      107 GTGAAATTTGATGCTCGCTCAATGACAGCTTTTGTCTCTCGAATCTCTAAACAGCCCT 166
Qy      41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60
Db      167 TCCCTTCAAGAGAAACTGAAAGTCTCTCAAAAGCTGCATGATTTGCCCTTTACCTCTCGTG 226
Qy      61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTyrGluThr 80
Db      227 TTTGCAAGTCTTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTTGAAGTGGGAACG 286
Qy      81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db      287 AGAATTTGCTCAGTTAGTTCACTAATGCAATGATATTAATCAAGTCTCACGGGAAA 346
Qy      100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db      347 GGAATGACACGCGAAGAGAAATGAGATTTCAAGAAGTCTTTATGGAACACATGAGCAAC 406
Qy      120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db      407 ATGGAGAAGAGAAATCCAGCATATTTTATAGCATGGAAGCCCACTCATGGACACAGAGCAT 466
Qy      140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db      467 TTCAAAATTTTCAGCATGACCACTGATCAAAAGATTTAATGACATTTCTTCGACGTAAGT 526
Qy      160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleGlyAspIleSerLysSerLeu 179
Db      527 ACCTTGTCTTCTCAGTCCAGGGACATGGGAATGCAATAGATGAAATCTCCAAAGTCTCTTA 586
Qy      180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db      587 ATAAAGTTTGAATACCACTTGTGATTTTCAGCTCAACATAGAAATCTGAAATGGCAAA 646
Qy      200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db      647 ATCCAGAGAAATACCTTCAAAACACACAGAGAAATCAGTAATATTAGAGGCGGTGTAC 706
Qy      220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
Db      707 AATGTATCAGCAGAAATTTATGGCTATGAAAGAGAACAAAGTGCATTTTGGAAACAGGAAATA 766
Qy      240 LysGlyGluMetLysLeuAsnIleThrAsnAspLeuArgLysLeuGluGluAspTyrGlu 259
Db      767 AAAGGAGAGTGAAGTGAATGAAATGCAATAGATGAAATCTTCAGACTGAAAGATTGGGAA 826
Qy      260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db      827 CATTCACAGACCTTGAGAAATATACCTTTAATTCAGGT 865

BD000149 1347 bp DNA linear PAT 31-JAN-2002
LOCUS

```



```

DEFINITION Antihuman scavenger receptor antibody.
ACCESSION BD000149
VERSION BD000149.1 GI:18623228
KEYWORDS JP 2000312595-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1347)
AUTHORS Kodama,T., Matsumoto,A. and Suzuki,H.
TITLE Antihuman scavenger receptor antibody
JOURNAL Patent: JP 2000312595-A 2 14-NOV-2000;
CHUGAI PHARM CO LTD
OS Homo sapiens (human)
PD JP 2000312595-A/2
PD 14-NOV-2000
PF 14-APR-2000 JP 2000114233
PR
PI TATSUHIKO KODAMA, AKIYO MATSUMOTO, HIROSHI SUZUKI PC
C12N15/09, C07K16/28, C12N5/10, C12P21/08// (C12N5/10, C12R1.91), PC
(C12P21/08, C12R1.91), C12N15/00, C12N5/00, (C12N5/00, C12R1.91) CC
FT Key Location/Qualifiers
FT source 1..1301
FT /organism="Homo sapiens (human)".
FEATURES
source
LOCATION/Qualifiers
1..1347
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 3.64e-65 Length: 1347
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 6 Gaps: 1
US-10-618-570-2 (1-400) x BD000149 (1-1347)
Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 47 ATGGAGCAGTGGGATCACTTTTCAAAATCAACAGGAGGACACTGATAGTCTCCGAATCT 106
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 107 GTGAAATTGATGCTCGCTCAACAGCTTGTCTTCTCCGAATCTCTTAAACAGCCCT 166
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
Db 167 TCCCTTCAAGAGAACTGAGTCTTCAAACTGCACTGATGTCCTTACTCTCGTG 226
Qy 61 PheValValLeuValProIleGlyLeuValAlaGlnLeuLeuLysTrpGluThr 80
Db 227 TTTCAGTTCTCATCTCTCTATTGGAATAGTGGCAGCTCAACTCTCTGAAGTGGAAACG 286
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 287 AAGAAATTCCTCAGTTAGTCTCACTAAATGCAATGATTAACCTCAAGTCTCACCGGAAA 346
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 347 GGAATGACAGGAGAGAGAAATGAGATTTCAGAGAGTCTTTATGGAACACATGAGCAAC 406
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGlnAlaAsnLeuLeuAspAlaLysAsn 139
Db 407 ATGGAGAGAGAAATCCACATATTTTAGACATGGAAGCCAACCTCATGGACACAGAGCAT 466
Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 467 TTCCAAATTTTCAGCACTGATCAAAAGATTTTATGACATTTCTTCTGCAGTAAAGT 526

```

```

Qy 160 SerLeuLeuSerSerIleGlnGlnHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
Db 527 ACCTTGTGTTTCTCAGTCCAGGAGCATGGGAATCAATAGATGAATCTCCAAGTCTTTA 586
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGlnThrLeuAsnGlyArg 199
Db 587 ATAAGTTTGAATACCAATTCCTTGTGATTCGAGTCAACATAGAAATCTGAAATGGCAAA 646
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 647 ATCCAAGAGATACCTTTCAAAACACAGAGGAAATCAGTAAATTAGAGGCGGTGTTTAC 706
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGlnLysGlnValTyrLeuGluGlnGluIle 239
Db 707 AATGTATCAGCAGAAATTTATGGCTATGAAAGAGAACAGTGCATTTTGGAAACAGAAATA 766
Qy 240 LysGlyGluMetLysLeuLeuAsnIleThrAsnAspLeuArgLysLeuAspTrpGlu 259
Db 767 AAGGAGAGTGAAGTACTGAAATCAATCACTAATGATCTCAGACTGAAAGATTGGGAA 826
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 827 CATCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 865

```

```

RESULT 11
LOCUS CQ722728 1367 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 8662 from Patent WO02068579.
ACCESSION CQ722728
VERSION CQ722728.1 GI:42283585
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 8662 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
Location/Qualifiers
source 1..1367
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 3.69e-65 Length: 1367
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 6 Gaps: 1
US-10-618-570-2 (1-400) x CQ722728 (1-1367)
Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 67 ATGGAGCAGTGGGATCACTTTCAAAATCAACAGGAGGACACTGATAGTCTCCGAATCT 126
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 127 GTGAAATTTGATGCTCGCTCAATGACAGCTTGTCTTCTCCGAATCTCTTAAACAGCCCT 186
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
Db 187 TCCCTTCAAGAGAACTGAGTCTTCAAAAGTCTCACTGATGTCCTTTACCTCTCGTG 246

```

Qy	61	PheValValLeuValProIleIleGlyIleValIalaIaGlnLeuLeuLysTyrGluThr	80
Db	247	TTTTGCAGTTCTCATCCCTCTCATTTGGAAATAGTGGCAGCTCAACTCTCTGAAGTGGGAACG	306
Qy	81	LysAsnCybThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys	99
Db	307	NAGAAITGCTCAGTTAGTTTCAACTAATGCAATGATATTAATCTMAAGTCTCAGCGGAATA	366
Qy	100	GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn	119
Db	367	GGAAATGACAGCGAAGAGAAATGAGATTTCAAGAAGTCTTTATGGAAACAATGAGCAAC	426
Qy	120	MetGluSerArgIleGlnTyrLeuSerSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn	139
Db	427	ATGGAGAAGAGAATCCAGCATATTTTACATGGAAGCCAACTCTGGGACACAGAGCAT	486
Qy	140	PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn	159
Db	487	TTCCAAAATTTTCAAGCATGCAACTGATCAAGATTTTAATGACATTTCTCTGAGCTTAAGT	546
Qy	160	SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu	179
Db	547	ACCTTGTCTTCATGCTCCAGGACATGGGAATGCAATAGATGAAATCTCCAAGTCTCTTA	606
Qy	180	ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg	199
Db	607	ATAAAGTTTGAATACCACTATGCTGTGATTCGACGTCAACATAGAAATCTGAATGGCAAA	666
Qy	200	ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr	219
Db	667	ATCCAAGAGAATACCTTCAACCAACAGAGAAATCAGTAAATTTAGAGGAGCGTGTTCAC	726
Qy	220	AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle	239
Db	727	AATGTATCAGCAGAAATATGCTGTATGAAGAAGAACAAAGTGCAATTTGGAAACAGGAATA	786
Qy	240	LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTyrGlu	259
Db	787	AAAGGAGAGAGTGAAGACTGATGAATAACATCACTAATGATCTCAGACTGAAGATGGGAA	846
Qy	260	HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly	272
Db	847	CATTCTCAGACCTTGAGAAATATCACTTTAATTCAGGT	885

Qy	1	MetAlaGlnTrpAspAspPheProAspGlnGlnIuAspThrAspSerCysThrGluSer	20
Ds	67	ATGAGCAGTGGGATCATCTTCACAATCAACAGGAGGACACTGATGACTGCTCCGAACTCT	126
Qy	21	VallysPheAspAlaAArgSerValThrAlaIeuLeuProHisProLysAsnGlyPro	40
Ds	127	GTGAATTTGATGCTCGCTCAATGACAGCTTTGCTCTCCGAATCTCAAAAACAGCCCT	186
Qy	41	ThrLeuGlnGluArgMetLysSerTyLysThrAlaLeuIleThrLeuTyLeuIleVal	60
Ds	187	TCCCTTCAAGAGAACTGAAAGTCTCTCAAGCTGCACTGATGGCCCTTTACCTCTCGTG	246
Qy	61	PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTrpGluThr	80
Ds	247	TTTGCACTTCTATCCCTCTCATTTGGAAATAGTGGCAGCTCAACTCTCTGAAGTGGAAACG	306
Qy	81	LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys	99
Ds	307	AAGAATTTGCTCAGTTAGTTCAACTTAATGCAAAATGATATACTCAAGCTCTCACGGGAAA	366
Qy	100	GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn	119
Ds	367	GGAAATCAGACGAGAGGAAATGAGATTTCAAGAGCTTTATGGAAACACATGAGCAAC	426
Qy	120	MetGluSerArgIleGlnTyLeuSerAspAsnGluAlaLeuLeuLeuAspAlaLysAsn	139
Ds	427	ATGAGAAAGAGAAATCCAGCATATTTTAGACATATGGAAGCAACCTCATGGAACACAGAGCAT	486
Qy	140	PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn	159
Ds	487	TTCAAAATTTTCAGCATGACATGATCAAAAGATTTAATGACATCTTCTGACAGCTAAGT	546
Qy	160	SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu	179
Ds	547	ACCTTGTTTCTCAGTCCAGGCAGCATGGGAATGCAATAGATGAAATCTCCAAGTCTCTTA	606
Qy	180	ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg	199
Ds	607	ATAAGTTTGAATACCACTGCTGATTTGTCAGCTCAACATAGAAAAATCTGAAATGGCAA	666
Qy	200	ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysIleGluGluArgIleTyr	219
Ds	667	ATCCAAGAGAAATACCTTCAAAACAAACAGAGAAATCAGTAAATAGAGGAGCGTGTTTAC	726
Qy	220	AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyLeuGluGlnGluIle	239
Ds	727	ANTGTATCAGCAGAAATATGGCTATGAAAGAGAGACAACTGCAATTTGGAAACAGGAATA	786
Qy	240	LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu	259
Ds	787	AAAGGAGAGTGAAGTACTGAAATCAATCACTAATGATCTCAGACTGAAAGATTTGGAA	846
Qy	260	HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly	272
Ds	847	CATTCTCAGACCTTGAGAAATATCACTTTAATTCAGGT	885
RESULT 13			
LOCUS AR270697 1367 bp DNA linear PAT 10-APR-2003			
DEFINITION Sequence 1260 from patent US 6500938.			
ACCESSION AR270697			
VERSION AR270697.1 GI:29701931			
KEYWORDS Unknown.			
SOURCE			

Alignment Scores:		
Pred. No.:	3,696-65	Length:
Score:	1042.50	Matches:
Percent Similarity:	87.2%	Conservative:
Best Local Similarity:	72.2%	Mismatches:
Query Match:	50.6%	Indels:
DB:	6	Gaps:
US-10-618-570-2 (1-400) x I72357 (1-1367)		

```

FEATURES             source
  1..1367             Location/Qualifiers
                        /organism="unknown"
                        /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.:           3.69e-65      Length:      1367
Score:               1042.50      Matches:    197
Percent Similarity:  87.2%        Conservative: 41
Best Local Similarity: 72.2%      Mismatches:  34
Query Match:        50.6%        Indels:     1
DB:                 8            Gaps:         1

US-10-618-570-2 (1-400) x AR270697 (1-1367)

Qy 1 MetAlaGlnTrpAspAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
Db 67 ATGGAGCAGTGGGATCACTTTACATCAATCAACAGGAGGACATGATAGCTTCGGAATCT 126
Qy 21 VallysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 127 GTGAAATTTGATGCTCGCTCAATGACAGCTTTGCTTCCTCGGAATCTTAAACACAGCCCT 186
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
Db 187 TCCCTTCAAGAGAACTCAAGTCTTCAAGCTGCACTGATTCGCTTTTACCTCTCGTG 246
Qy 61 PheValValLeuValProIleGlyLeuValAlaAlaGlnLeuLeuLysTrpGluThr 80
Db 247 TTTGAGTTTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTGAACTGGAACG 306
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 307 AAGATTGCTCAGTTAGTTCAACTAATGCAATGATATACTCAAGTCTCACGGGAAA 366
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 367 GGAATGACAGCGAAGAGAAATGAGATTTCAAGAAGTCTTTATGGAACACATGAGCAAC 426
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 427 ATGAGAAGAGAAATCCAGCATATTTTACATGATGAAGCAACCTCATGGACACAGAGCAT 486
Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 487 TTCCAAATTTTCAGCATGACACTGATCAAGATTTAATGACATCTTCTCGACTAAGT 546
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleLeuGlyAspIleSerLysSerLeu 179
Db 547 ACCTTGTCTTCTCAGTCCAGGACATGGGAATGCAATAGATGAATCTTCCAACTCTTA 606
Qy 180 ValGlyLeuAsnThrThrValLeuAspIleGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 607 ATAAAGTTGAATACCACTCTGATTTGAGTTCAGCTCAACATAGAAAATCTGAATGGCAAA 666
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 667 ATCCAGAGATACCTTCAACACACAGAGAAATCAGTAATTAATGAGGACGCTTTAC 726
Qy 220 AsnAlaSerAlaGluLeuLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluLe 239
Db 727 AATGTATCAGCAGAAATTTATGCTATGAAAGAAAGCAAGTGCAATTTGGAACAGGAATA 786
Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
Db 787 AARGAGAGAGTGAAGTACTGAATAACATCATCTAATGATCTCAGACTGAAAGATTGGAA 846
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 847 CATTTCTAGACCTTGAGAAATATCATCTTTAATTCAGGT 885

RESULT 14

```

```

HUMPHSR2             1367 bp      mRNA      linear      PRI 18-DEC-2002
LOCUS                Homo sapiens phsR2 mRNA for scavenger receptor type II, complete
                        cds.
ACCESSION            D90188
VERSION              D90188.1 GI:219991
KEYWORDS
SOURCE               Homo sapiens (human)
ORGANISM
AUTHORS
REFERENCE
  1 (bases 1 to 1367)
  Matsumoto,A., Naito,M., Itakura,H., Ikemoto,S., Aaoka,H.,
  Hayakawa,I., Kanamori,H., Aburatani,H., Takaku,F., Suzuki,H.,
  Kobari,Y., Miyai,T., Takahashi,K., Cohen,H.E., Wydro,R.,
  Houman,E.D. and Kodama,T.
  Human macrophage scavenger receptors: primary structure,
  expression, and localization in atherosclerotic lesions
  Proc. Natl. Acad. Sci. U.S.A. 87 (23), 9133-9137 (1990)
  2251254
COMMENT
  These data kindly submitted in computer readable form by: Akiyo
  Matsumoto
  The National Institute of Health and Nutrition
  1-23-1 Toyama
  Shinjuku-ku, Tokyo 162
  Phone: 81-3-3203-5725
  Fax: 81-3-3207-3520
  Location/Qualifiers
    1..1367
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /note="monocytic cell line THP-1 (4 days after phorbol
      ester treatment)"
    gene
      1..1367
      /gene="phsR2"
      67..1143
      /gene="phsR2"
      /codon_start=1
      /product="scavenger receptor type II"
      /protein_id="BAA14209.1"
      /db_xref="GI:219992"
      /translation="MEQWDFHMQOEDTDCSRVSKFDARSMTALLPPNKNPSLOE
      KLSFKAAALYLLVPAVLPIIGIIVAOLKWKETKNCVSVSTNANDITQSLTKGN
      DSEEMRPFQVFMHSMNMEKRIQHLIDMEANLMDTEHFQNFSTTDQRFNDILLQLS
      TLFSSVGHGNAIDEISKLSISLNTLLDLDQLNENLNGIKIQENTFKQBEISKLER
      VYNVSAIIMAKBEEQVHLEQIEIKGEVKNNTNDLRLKDWHSQTLRNTLIQPPG
      PPGKGDGRGTGSGPRGPGPIGPPGLKGDRAIGPPGSRGLPGYAGRPNGSGPKGQ
      KBEKSGNTLRPVQLTDHIRAGPS"
ORIGIN
Alignment Scores:
Pred. No.:           3.69e-65      Length:      1367
Score:               1042.50      Matches:    197
Percent Similarity:  87.2%        Conservative: 41
Best Local Similarity: 72.2%      Mismatches:  34
Query Match:        50.6%        Indels:     1
DB:                 8            Gaps:         1

US-10-618-570-2 (1-400) x HUMPHSR2 (1-1367)

Qy 1 MetAlaGlnTrpAspAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
Db 67 ATGGAGCAGTGGGATCACTTTACATCAATCAACAGGAGGACATGATAGCTTCGGAATCT 126
Qy 21 VallysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 127 GTGAAATTTGATGCTCGCTCAATGACAGCTTTGCTTCCTCGGAATCTTAAACACAGCCCT 186
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60

```

```
Db 187 TCCCTTCAAGAGAACTGAAGTCCTTCAAGCTGCACCTGATTGCCCTTTACCTCCTCGTG 246
Qy 61 PheValValLeuValProIleGlyLeValAlaAlaGlnLeuLeuValTrpGluThr 80
Db 247 TTTCAGATTCTATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCCTGGAATGGGAACG 306
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 307 AAGAATTGCTCAGTTAGTTCAACTAATGCAATATGATATACTCAAGATCTCACGGGAAA 366
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 367 GGAATGACAGCAGAGAGAAATGAGATTTCAAGAAGTCTTTATGGAACACATGAGCAAC 426
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGlnAlaAsnLeuLeuAspAlaLysAsn 139
Db 427 ATGGAAGAAGAAATCCACATATTTAGACATGGAAGCCAACTCATGGACACAGAGCAT 486
Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 487 TTCCAATAATTTTCAGCATGACAACTGATCAAGATTTTAATGACATTTCTTCGACGTAAGT 546
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIlelleGlyAspIleSerLysSerLeu 179
Db 547 ACCTTGTGTTCTCAGTCCAGGACATGGGAATGCAATAGATGAATCTCCAAGTCTCTTA 606
Qy 180 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 607 ATAAAGTTGAATACACATTCCTGATTGTCAGCTCAACATAGAAATCTGAATGGCAAA 666
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 667 ATCCAAGAGATACCTTCAACACACAGAGAAATCAGTAATTTAGAGGCGGTGTTCAC 726
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
Db 727 AATGATATCAGCAGAAATATGCTGATGAAAGAAAGAACAAAGTGCAATTTGGAAACGAAATA 786
Qy 240 LysGlyGluMetLysLeuLeuAsnIleThrAsnAspLeuArgLeuLysAspTyrGlu 259
Db 787 AAGAGAGAAAGTGAAGTACTGAATAACATCACTAATGATCTCAGACTGAAAGATGGGAA 846
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 847 CATCTCAGACCTTGAGAAATATCACITTAATTCAGGT 885

RESULT 15
E03267 LOCUS E03267 2028 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA sequence coding for human scavenger receptor, type 1.
ACCESSION E03267
VERSION E03267.1 GI:2171484
KEYWORDS JP 1991290184-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2028)
AUTHORS Matsumoto,A. and Kodama,T.
TITLE CELL OF SCAVENGER RECEPTOR-PRODUCING ANIMAL
JOURNAL Patent: JP 1991290184-A 1 19-DEC-1991;
CHUGAI PHARMACEUT CO LTD
OS Homo sapiens (human)
PN JP 1991290184-A/1
PD 19-DEC-1991
PF 06-APR-1990 JP 1990090274
PI MATSUMOTO AKIYO, KODAMA TATSUHIKO
PC C12N5/10.C12N15/12.C12Q1/06;
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
```

```
CC *source: cell type=monocyte leukemic cells;
CC *source: cell_line=THP-1;
FH Key Location/Qualifiers
FH 5'UTR 1..46
FH CDS 47..1402
FT /product='human scavenger receptor, type 1' FT
FT 1403..2028.
FEATURES
source
3'UTR Location/Qualifiers
1..2028
/organism="Homo sapiens"
/db_type="genomic RNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 5.58e-65 Length: 2028
Pred. No.: 1042.50 Matches: 197
Score: 1042.50
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 6 Gaps: 1
US-10-618-570-2 (1-400) x E03267 (1-2028)
```

```
Qy 1 MetAlaGlnTrpAspAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 47 ATGGAGCAGTGGGATCCTTTTCACATCAACAGAGACACTGATAGCTGCTCGAATCT 106
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 107 GTGAATTTGATGCTCGCTCAATGACAGAGCTTTGCTTCTCCGAATCTTAAACACAGCCCT 166
Qy 41 ThrLeuGlnGluArgMetLysSerTyrThrAlaLeuIleThrLeuTyrLeuIleVal 60
Db 167 TCCCTTCAAGAGAACTGAAGTCTCTCAAAAGCTGCACCTGATTGCCCTTTTACCTCTCGTG 226
Qy 61 PheValValLeuValProIlelleGlylleValAlaAlaGlnLeuLeuValTrpGluThr 80
Db 227 TTTCAGTTCTCATCCCTCTCATTTGGATAGTGGCAGCTCACTCTCTGAGTGGGAACG 286
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 287 AAGAAATTCCTCAGTTAGTTCAACTAATGCAATATGATATACTCAAGTCTCACGGGAAA 346
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 347 GSAATGACACGAGAGGAAATGAGATTTCAAGAAAGTCTTTATGGAACACATGAGCAAC 406
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGlnAlaAlaAsnLeuLeuAspAlaLysAsn 139
Db 407 ATGGAGAGAGATCCAGCATATTTTAGACATGGAGCCCACTCATGGACACAGAGCAT 466
Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 467 TTCCAAAATTTTCAGCATGACCACTGATCAAGATTTTAATGACATTTCTTCGACGTAAGT 526
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIlelleGlyAspIleSerLysSerLeu 179
Db 527 ACCTTGTGTTCTCAGTCCAGGACATGGGAATGCAATAGATGAAATCTCCAAGTCTCTTA 586
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 587 ATAAGTTTGAATACCATTTGTTGATTTCAGCTCAACATAGAAATCTGAATGGCAAA 646
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 647 ATCCAAGAGATACCTTCAACACACAGAGGAAATCAGTAATTTAGAGGAGCGGTGTTCAC 706
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
Db 707 AATGATATCAGCAGAAATTTATGCTATGAAAGAGAAACAAAGTGCAATTTTGGAAACAGGAATA 766
```

Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 767 AARGGAGAGTGAAAGTACTGAATACATCACTAATGATCTCAGACTGAAGATTGGAA 826
 Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 827 CATTCAGACCTTGAGAAATATCACTTTAATTCAGGT 865

Search completed: February 27, 2006, 10:41:09
 Job time : 6664 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 27, 2006, 09:14:37 ; Search time 5148 Seconds
(without alignments)
3635.360 Million cell updates/sec

Title: US-10-618-570-2

Perfect score: 2059

Sequence: 1 MAQWDDPQDQEDTDSCTES.....DWKATRVGINIFRLRTOKE 400

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWEB spool/US10618570/runat 27022006 063856 13092/app query.fasta_1
-DB=EST -QPWT=fastap -SUFFIX=rest -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORES=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZES=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs807
-USER=US10618570 @CGN 1 1 5315 @runat 27022006 063856 13092 -NCPU=6 -ICPU=3
-NO.MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*

Query	Match	Length	ID	Description
1: gb_est1:*	62.0	793	8	DN548862
2: gb_est2:*	55.4	949	6	CF409555
3: gb_est3:*	50.6	1033	10	AY413208
4: gb_est4:*	49.5	652	6	CB453082
5: gb_est5:*	48.7	1027	10	AY413209
6: gb_est6:*	45.9	978	7	CN643776
7: gb_est7:*	43.0	843	5	BX496563
8: gb_est8:*				
9: gb_est9:*				
10: gb_est10:*				
11: gb_est11:*				

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1276	62.0	793	8	DN548862 1407127 M
2	1140.5	55.4	949	6	CF409555 CH3#061 D
3	1042.5	50.6	1033	10	AY413208 Homo sapi
4	1020	49.5	652	6	CB453082 708007 MA
5	1002.5	48.7	1027	10	AY413209 Pan trogl
6	944.5	45.9	978	7	CN643776 ILLUMIGEN
7	885	43.0	843	5	BX496563 DKFZp779D

8	830	40.3	3635	4	AK028480
9	823	40.0	1030	10	AY413210
10	801	38.9	1509	4	AK089178
11	791.5	38.4	677	5	BW956649
12	768.5	37.3	768	8	DR004903
13	736.5	35.8	626	6	CB469554
14	727.5	35.3	651	8	DR004296
15	719.5	34.9	1093	7	CN643736
16	717	34.8	816	6	CB997441
17	714.5	34.7	665	6	CB472122
18	698.5	33.4	597	6	CB990950
19	678	32.9	559	6	CF251702
20	678	32.9	561	6	CD725634
21	678	32.9	575	5	BU447766
22	678	32.9	576	5	BU408647
23	678	32.9	582	5	BX269637
24	678	32.9	586	6	CF251883
25	678	32.9	588	6	CF252572
26	678	32.9	594	6	CF252815
27	674	32.7	581	6	CF251833
28	673	32.7	576	5	BU447569
29	670	32.5	570	5	BU409120
30	670	32.5	578	5	BU426481
31	670	32.5	612	5	BU422332
32	669.5	32.5	585	5	BW958294
33	665	32.3	594	6	CF252175
34	664	32.2	538	3	BM489973
35	657	31.9	473	3	BQ038085
36	657	31.9	693	1	AW744595
37	654	31.8	589	6	CF251795
38	649	31.5	453	7	CK608106
39	649	31.5	458	7	CK614089
40	647	31.4	456	7	CK614219
41	647	31.4	721	6	CF745486
42	645	31.3	460	7	CK615416
43	641	31.1	460	7	CK608142
44	641	31.1	465	7	CK614950
45	637	30.9	593	5	BU418594

ALIGNMENTS

RESULT 1

DN548862

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

DN548862 1407127 MARC 7BOV Bos taurus CDNA 5', mRNA sequence. EST 11-MAR-2005

DN548862 1407127 MARC 7BOV Bos taurus CDNA 5', mRNA sequence.

DN548862.1 GI:61013550

EST.

Bos taurus (cow)

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;

Pecora; Bovidae; Bovinae; Bos.

Smith, T.P.L., Roberts, A.J., Echtenkamp, S.E., Chitko-McKown, C.G.,

Wray, J.E. and Keefe, J.W.

A second set of bovine ESTs from pooled-tissue normalized libraries

Unpublished (2003)

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called with phred v0.020425.c and

trimmed with the aid of the trim_alt option. Vector identified with

cross match v0.990329.

Plate: QOP8005 row: C column: 19

Seq primer: GTAATACGACTACTATAGG.

Location/Qualifiers

1..793

/organism="Bos taurus"

FEATURES

source

```

/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 7B0V"
/notes="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including ovary, hindbrain, uterus, and day-30 whole
embryos."

ORIGIN
Alignment Scores:
Pred. No.: 1,248-138 Length: 793
Score: 1276.00 Matches: 252
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 62.0% Indels: 0
DB: 8 Gaps: 0

US-10-618-570-2 (1-400) x DN548862 (1-793)

Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db ATGGCACAGTGGATGACTTTCCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCT 95
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db GTGAAGTTCGATGCTCGCTCAGTGACAGCTTTGGCTTCTCCCTCCCTAAATAATGGCCCA 155
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
Db ACTCTTCAGAGAGATGAGTCTTTATTAACCTGACCTGATCACCCTTATCTCATGTG 215
Qy 61 PheValValLeuValProIleGlyIleValAlaAlaGlnLeuLeuTyrTrpGluThr 80
Db TTTGTAGTTCCTGTGCCCATCATTTGGCATAGTGCCAGCTCAGCTCCTGAAATGGGAAACG 275
Qy 81 LysAsnCysThrValGlySerValAsnAlaAspLysSerProSerProGluGlyLysGly 100
Db AAGAAATTCACGGTTGGCTCAGTTAATGCAGATATATCTCAAGTCCGGAAGGCCAAGGA 335
Qy 101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120
Db AATGCGATGAGATGAATGAGATTCGAGAGAGCTGTGATGGAACGATGAGCAACATG 395
Qy 121 GluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPhe 140
Db GAAAGCAGATCCAGTATCTTTCAGATTAATGAAGCCAATCTCTAGATGCTTAAGAAATTC 455
Qy 141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
Db CAAATTTTCAGCATAAACAACTGATCAAAAGATTTTAATGATGTCTTTTCCAGCTAAATTC 515
Qy 161 LeuLeuSerSerIleGlnGlnHisGluAsnIleLeuGlyAspIleSerLysSerLeuVal 180
Db TTACTTTCTCCATCCAGAACATGAGAAATATCATAGGGGATATCTCCAACTCATAGTA 575
Qy 181 GlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal 200
Db GGTCTGAACACACAGTACTTGTATTTGCGATTCAGTATTGAAACACTGAATGGCAGAGTC 635
Qy 201 GlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyrAsn 220
Db CAAGAGAAATGCAATTTAAACAAAGAGAGAGATGCGTAAATTTAGAGGAGCGGTATATCAAT 695
Qy 221 AlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluLys 240
Db GCATCAGCAGAAATTAAGTCTCTAGATGAAACAAAGTATATTTTGGACAGGAATATAA 755
Qy 241 GlyGluMetLysLeuLeuAsnAsnIleThrAsnAsp 252
Db GGGGAAATGAACATGTTTGAATAATATATCACTAATGAT 791

```

RESULT 2
CF409555
LOCUS
DEFINITION
CF409555
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

CF409555 949 bp mRNA linear EST 02-SEP-2003
CH3#061_D06MF Canine heart normalized cDNA Library in pBluescript
Canis familiaris cDNA clone CH3#061_D06 5', mRNA sequence.
CF409555
CF409555.1 GI:34410801
EST.
Canis familiaris (dog)
Canis familiaris
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.
1 (bases 1 to 949)
Yi, Y., Desai, R., Olarte, M., Henthorn, P. and George A.L.
Expressed sequence tags from Canine heart
Unpublished (2003)
Other ESTs: CH3#061_D06MR
Contact: George AL_
Division of Genetic Medicine
Vanderbilt University
529 Light Hall, 2215 Garland Avenue, Nashville, TN 37232-0275, USA
Tel: 615 936 2660
Fax: 615 936 2661
Email: al.george@vanderbilt.edu
Insert Length: 2287 Std Error: 0.00
Seq primer: MF: GTTTCCTCCAGTCACGACGTG
High quality sequence start: 169
High quality sequence stop: 784.
Location/Qualifiers
1..949
/organism="Canis familiaris"
/mol_type="mRNA"
/db_xref="taxon:9615"
/clone="CH3#061_D06"
/tissue_type="heart"
/cell_type="heart"
/dev_stage="mixed developmental stages (adult, 30 day - 40 day fetal)"
/clone_lib="Canine heart normalized cDNA Library in pBluescript"
/notes="Organ: heart; Vector: pBluescript; Site 1: 5' of vector NotI; Site 2: 3' of vector EcoRI; Tissue source: dog heart (adult, 30 day - 40 day fetal), right and left atria and ventricle. Dog breed - mixed (beagle, German shepherd, pointer, Irish setter). Library construction: oligo-dT primed"

Alignment Scores:
Pred. No.: 1,328-122 Length: 949
Score: 1140.50 Matches: 235
Percent Similarity: 96.0% Conservative: 3
Best Local Similarity: 94.8% Mismatches: 6
Query Match: 55.4% Indels: 4
DB: 6 Gaps: 3

US-10-618-570-2 (1-400) x CF409555 (1-949)

Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db ATGGCACAGTGGATGACTTTCCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCT 264
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db GTGAAGTTCGATGCTCGCTCAGTGACAGCTTTGGCTTCTCCCTCCCTAAATAATGGCCCA 324
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
Db ACTCTTCAGAGAGATGAGTCTTTATTAACCTGACCTGATCACCCTTATCTCATGTG 384
Qy 61 PheValValLeuValProIleGlyIleValAlaAlaGlnLeuLeuTyrTrpGluThr 80


```

Db      385  TTTGTAGTTCTCGTCCCATCTGGCATAGTGGCAGCTCAGCTCCTGAAATGGGAACG 444
Qy      81  LysAnCysThrValGlySerValAsnAlaAspIleSerProSerProGluGlyIleGly 100
Db      445  AAGAATTGACGGTGGCTGATTAATGACATATATCTCCAACTCCGGAAGGCAAGGA 504
Qy      101  AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120
Db      505  AATGGCAGTGAAGATGAATGAGATTTTCGAGAAGCTGTGATGGAAACGATGAGCAACATG 564
Qy      121  GluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPhe 140
Db      565  GAAAGCAGATCCAGTATCTTTTCAGATAATGAAGCCCAATCTCTAGATGCTAAGAAATTC 624
Qy      141  GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
Db      625  CAAAATTTTCAGCAATCAACACTGATCAAGATTTAATGATGTTCTTTTCCAGCTAAATTC 684
Qy      161  LeuLeuSerSerIleGlnGluHicGluAsnIleIleGlyAspIleSerLysSerLeuVal 180
Db      685  TTACTTTCTCCATCCAGGAACATGAGAAATATCATAGGGGATATCTCCAACTCATTAGTA 744
Qy      181  GlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal 200
Db      745  GGTCTGACACACACAGTACTTGTATTTGAGTTTCACTTATTTGAAACACTGAATGGCAGATC 804
Qy      201  GlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyrAsn 220
Db      805  CAAGAGAATGCATTTAAACCAACAGAGAGATGCGTAAATTTAGAGGAGCGGTATATACAAT 864
Qy      221  AlaSerAlaGlu-tleYsSerLeuAspGluLys---GlnValTyrLeu---GluGlnG 238
Db      865  GCATCAGCGAATAATTAAGTCTCTAGATGGAACCAACCAAGTATATTTTGGGAACAGGA 924
Qy      238  uileLys---GlyGluMetLys 244
Db      925  AATTAAAGGGGGAATGGAA 946

RESULT 3
AY413208
LOCUS      1033 bp      DNA      linear      GSS 17-DEC-2003
DEFINITION Homo sapiens MSRI gene, VIRTUAL TRANSCRIPT, partial sequence,
GENOMIC SURVEY SEQUENCE.
ACCESSION AY413208
VERSION   AY413208.1 GI:39769173
KEYWORDS GSS.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 1033)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
          Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
          Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
          Adams,M.D. and Cargill,M.
          Inferring nonneutral evolution from human-chimp-mouse orthologous
          gene trios
JOURNAL   Science 302 (5652), 1960-1963 (2003)
PUBMED    14671302
REFERENCE 2 (bases 1 to 1033)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
          Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
          Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
          Adams,M.D. and Cargill,M.
          Direct Submission
JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
          Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
          them based on alignment.
FEATURES  Location/Qualifiers
           source          1..1033
           /organism="Homo sapiens"

```

```

/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1033
/gene="MSRI"
/locus_tag="HCM4814"

ORIGIN

Alignment Scores:
Pred. No.:      4.82e-111      Length:      1033
Score:          1042.50        Matches:    197
Percent Similarity: 87.2%      Conservative: 41
Best Local Similarity: 72.2%    Mismatches: 34
Query Match:    50.6%          Indels:     1
DB:             10             Gaps:       1

US-10-618-570-2 (1-400) x AY413208 (1-1033)

Qy      1  MetAlaGlnTrpAspAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
Db      1  ATGGAGCAGTGGGATCACTTTCACATCAACAGAGGACACTGATAGCTGCTCCGAATCT 60
Qy      21  ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db      61  GTGAATTTGATGCTCGCTCAATGACAGCTTTGCTTCTCCGATCTTAAACACAGCCCT 120
Qy      41  ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuIleVal 60
Db      121  TCCCTTCAAGAGAACTGAAGTCTCTCAATGACAGCTTTGCTTCTCCGATCTTAAACACAGCCCT 180
Qy      61  PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTyrGluThr 80
Db      181  TTTGCACTTCTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTGGAATGGGAACG 240
Qy      81  LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db      241  AAGNATTTGCTCAGTTAGTTTCACTAATGCAATGATATATACTCAAGTCTCACGGGAAA 300
Qy      100  GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db      301  GGAATGACAGCGAAGAGAAATGAGATTTCAAGAACTCTTTATGGAACACATGAGCAAC 360
Qy      120  MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db      361  ATGGAGAAAGAGAAATCCAGCATATTTTAGACATGGAAGCCCAACTCATGGACACAGAGAT 420
Qy      140  PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db      421  TTTCCAAATTTTCAGTCATGACACTGATCAAGATTTAATGACATCTTCTGCAGCTAAGT 480
Qy      160  SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
Db      481  ACCTTGTCTTCTCAGTCAGGAGACATGGGAATGCAATAGATGAATCTCCAGTCTCTTA 540
Qy      180  ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db      541  ATAAGTTTGAATACCACTTGTGATTTTCAGCTCAACATAGAAAATCTGAATGGCAAA 600
Qy      200  ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db      601  ATCCAGAGATACCTTCAAAACACAGAGAAATCATGTAATTTAGAGGAGCGTGTATAC 660
Qy      220  AsnAlaSerAlaGluIleLysSerLysGlnLysGlnValTyrLeuGluGlnGluIle 239
Db      661  AATGATATCAGCAGAAATATTATGGCTATGAAAGAGAAACAACTGATTTTGGAAACAGGAATA 720
Qy      240  LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTyrGlu 259
Db      721  AAAGGAGAGTGAAGAGTACTGAAATACATCACTAATGATCTCAGACTGAAGATGGGAA 780
Qy      260  HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db      781  CATTTCTCAGACCTTGAAGAAATATACACTTTTAATTCAAGGT 819

```

```

RESULT 4
CB453082          652 bp      mRNA      linear      EST 26-MAR-2003
LOCUS             708007 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION        CB453082
ACCESSION         CB453082
VERSION           CB453082.1 GI:29259464
KEYWORDS          EST.
SOURCE            Bos taurus (cow)
ORGANISM          Bos taurus
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
                  Pecora; Bovidae; Bovinae; Bos.
REFERENCE          1 (bases 1 to 652)
AUTHORS           Smith,T.P.L., Roberts,A.J., Echterkamp,S.E., Chitko-McKown,C.G.,
                  Wray,J.E. and Keele,J.W.
TITLE             A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL            Unpublished (2003)
COMMENT           Contact: Smith TPL
                  USDA, ARS, US Meat Animal Research Center
                  PO Box 166, Clay Center, NE 68933-0166, USA
                  Tel: 402 762 4366
                  Fax: 402 762 4390
                  Email: smith@mail.marc.usda.gov
                  Single pass sequencing. Bases called with phred v0.020425.c and
                  trimmed with the aid of the trim_alt option. Vector identified with
                  cross_match v0.990329.
                  Plate: PQ8065 row: M column: 21
                  Seq primer: GTAATGACCTCCTACTATAGG.
FEATURES          Location/Qualifiers
                    1..652
                        /organism="Bos taurus"
                        /mol_type="mRNA"
                        /db_xref="taxon:9913"
                        /tissue_type="pooled"
                        /lab_host="DH10B"
                        /clone_lib="MARC 6BOV"
                        /note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
                        Library made with RNA pooled from multiple tissues
                        including liver, lung, hypothalamus, pituitary, and
                        placenta/endometrium."
ORIGIN
Alignment Scores:
Pred. No.:        1.05e-108      Length:      652
Score:            1020.00         Matches:     207
Percent Similarity: 81.8%         Conservaive: 0
Best Local Similarity: 81.8%      Mismatches: 0
Query Match:      49.5%          Indels:      46
DB:               6              Gaps:        1

US-10-618-570-2 (1-400) x CB453082 (1-652)

Qy      1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db      32 ATGGCACAGTGGGATGACCTTCTCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCT 91
Qy      21 ValLysPheAspAlaAArgSerValThrAlaLeuLeuProProHisProLysAsnGlyPro 40
Db      92 GTGAAGTTTCGATGCTCGCTCAGTGACAGCTTTGCTTCTCCCATCTCTAAATAATGGCCCA 151
Qy      41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuLeuVal 60
Db      152 ACTCTTCAAGAGAGATGAGTCTATTAATACTGACATGATCACCCCTTATCTCATTTGTG 211
Qy      61 PheValValLeuValProIleGlyLeValAlaAlaGlnLeuLeuTyrGluThr 80
Db      212 TTGTAGTTCTCGTCCCATCATTTGGCATATGTCAGCTCAGCTCCTGAAATGGAAACG 271
Qy      81 LysAsnCysThrValGlySerValAsnAlaAspLysProSerProGluGlyLysGly 100
Db      272 AAGAATTGCACGGTTGGTTCAGTTAATGCAGATATATCTCCAGTCCGGAAGGCAAGGA 331
Qy      101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120

```

```

Db      332 AATGGCAGTGAAGATGAATGAGATTTTCGAGAAGCTGTGATGAACGATGAGCAACATG 391
Qy      121 GluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPhe 140
Db      391 -----
Qy      141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
Db      391 -----
Qy      161 LeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspLysSerLysLeuVal 180
Db      392 -----GAACATGAGATATCATAGGGGATATCTCCAAGTCATTAGTA 433
Qy      181 GlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal 200
Db      434 GGTCTGAACACCACTACTTGTGATTTGTCAGTTTCAGTATTGAAAACACTGAATGCGAGATC 493
Qy      201 GlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyrAsn 220
Db      494 CAAGAGATGCATTTAAACCAAGAGAGATCGTAAATTAGAGGAGCGTATATACAT 553
Qy      221 AlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIleLys 240
Db      554 GCATCAGCAGAGAAATTAAGTCTCTAGATGAAAAACAAGTATATTTGGAACAGGAATAAAA 613
Qy      241 GlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeu 253
Db      614 GGGGAATGAACCTGTTGAATATATATCACTAATGATCTG 652

RESULT 5
AY413209          1027 bp      DNA      linear      GSS 17-DEC-2003
LOCUS             Pan troglodytes MSR1 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION        genomic survey sequence.
ACCESSION         AY413209
VERSION           AY413209.1 GI:39769174
KEYWORDS          GSS.
SOURCE            Pan troglodytes (chimpanzee)
ORGANISM          Pan troglodytes
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
                  Homidae; Pan.
REFERENCE          1 (bases 1 to 1027)
AUTHORS           Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
                  Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
                  Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
                  Adams,M.D. and Cargill,M.
TITLE             Inferring nonneutral evolution from human-chimp-mouse orthologous
JOURNAL            Science 302 (5652), 1960-1963 (2003)
PUBMED            14671302
REFERENCE          2 (bases 1 to 1027)
AUTHORS           Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
                  Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
                  Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
                  Adams,M.D. and Cargill,M.
TITLE             Direct Submission
JOURNAL            Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
                  Rockville, MD 20850, USA
COMMENT           This sequence was made by sequencing genomic exons and ordering
                  them based on alignment.
FEATURES          Location/Qualifiers
                    1..1027
                        /organism="Pan troglodytes"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9598"
                        <1..>1027
                        /gene="MSR1"
                        /locus_tag="HCM4814"
ORIGIN

```

Alignment Scores:
 Pred. No.: 2,378-106 Length: 1027
 Score: 1002.50 Matches: 191
 Percent Similarity: 84.6% Conservative: 39
 Best Local Similarity: 70.2% Mismatches: 41
 Query Match: 48.7% Indels: 1
 DB: 10 Gaps: 1

US-10-618-570-2 (1-400) x AY413209 (1-1027)

Qy 1 MetAlaGlnTrpAspAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
 Db 1 ATGGAGCAGTGGGATCACTTTACAAATCAACAGAGGACACTAGTGTCTGNNATCT 60

Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
 Db 61 GTGAAATTTGATCGCTCGCTCAATGACAGCTTTGCTTCTCCGAAATCTTAAACACAGCCCT 120

Qy 41 ThrLeuGlnGluArgMetLysSerTyrlsThrAlaLeuLeuThrLeuTyrlsLeuVal 60
 Db 121 TCCCTTCAAGAGAACTGAAGTCTTCAAGCTGCACTGATTCCTTACCTCTCTGCTG 180

Qy 61 PheValValLeuValProIleGlyLeuValAlaAlaGlnLeuLeuLysTrpGluThr 80
 Db 181 TTTCAGTTCATCTCCCTCTCATTTGGAATAGTGGCAGCTCACTCTGAAATGGGAACG 240

Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
 Db 241 AAGAATTGCTCAGTTGGTTCAACTAATGCAATGATATACTCAAGTCTCACGGGAAA 300

Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
 Db 301 GGAAATGACACGGAAGAGGAATGAGATTTCAAGAAGTCTTTATGGACACATGAGCAAC 360

Qy 120 MetGluSerArgIleGlnTyrlsSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
 Db 361 ATGGAGAAGAAATCCACATATTTTCAGCATGGAAGCCAACTCATGGACACAGAGCAT 420

Qy 140 PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
 Db 421 TTCCAAATTTTCAGCATGACAACTGATCAAGATTTAATGACATCTTCTGCAGCTAAGT 480

Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIlelleGlyAspIleSerLysSerLeu 179
 Db 481 ACCTTGTCTTCTCAGTCNNNNNNATGGGAATGCAATAGATGAATCTCCAAGTCTCTTA 540

Qy 180 ValGlyLeuAsnThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
 Db 541 ATAAAGTTTGAATACCAATNNNTGATTTGAGCTCAACATAGAAAATCTGAATGGCAAA 600

Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyrl 219
 Db 601 ATCCAGAGATACCTTCAACNNNNNNNNGAATCAGTAATTTAGAGGCGTGTTCAC 660

Qy 220 AsnAlaSerAlaGluLeuLysSerLeuAspGlyLysGlnValTyrlLeuGluGlnGluLe 239
 Db 661 AATGTATCAGCAGAAATATGCTGATGAAGAGAACAAAGTGATTTTGAACAGCAATA 720

Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAsnTrpGlu 259
 Db 721 AAAGGAGAGTGAAGTACTGAATAACATCACTAATGATCTCAGACTGAAAGATTTGGAA 780

Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGln 271
 Db 781 CATCTCAGACTTGAGAAATATCACTTATTCAA 816

RESULT 6
 CN643776
 LOCUS
 DEFINITION
 5' similar to Bases 47 to 976 highly similar to human MSRI
 (Hs.436887), mRNA sequence.
 CN643776
 ACCESSION

VERSION
 EST.
 SOURCE
 ORGANISM
 Macaca mulatta (rhesus monkey)
 Macaca mulatta
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Cercopithecoidea; Cercopithecinae; Macaca.
 1 (bases 1 to 978)
 Magness,C.L., Fellin,P.C., Thomas,M.J., Korth,M.J., Agy,M.B.,
 Proll,S.C., Fitzgibbon,M., Scherer,C.A., Miner,D.G., Katze,M.G. and
 Iadonato,S.P.
 Analysis of the Macaca mulatta transcriptome and the sequence
 divergence between Macaca and human
 Genome Biol. 6 (7), R60 (2005)
 15998449
 COMMENT
 Contact: C. Magness
 Illumigen Biosciences Inc.
 2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
 Tel: 2063780400
 Fax: 2063780408
 Email: cmagness@illumigen.com
 Sequenced on 2004.01.16. 752 Q20 bases.
 PCR Primers
 FORWARD: CCCTCACTAAAGGGAACAAAA
 BACKWARD: CACTATAGCGCGAATGGGTA
 Insert Length: 978 Std Error: 0.00
 Plate: CL000089 row: C column: 03
 Seq primer: CCCTCACTAAAGGGAACAAAA
 POLYA=NO.

FEATURES
 source
 1..978
 /organism="Macaca mulatta"
 /mol_type="mRNA"
 /strain="Indian"
 /db_xref="taxon:9544"
 /clone="IBIUM:3182"
 /sex="female"
 /dev_stage="adult"
 /lab_host="E. coli SOLR"
 /clone_lib="Katze WMBR"
 /note="Organ: brain; Vector: Uni-ZAP XR; Site 1: EcoR I;
 Site 2: Xho I; Created from Stratagene ZAP-CDNA Synthesis
 kit (catalog #200400) and ZAP-CDNA Gigapack III Gold
 Cloning Kit (Catalog #200450)"

ORIGIN
 Alignment Scores:
 Pred. No.: 1,428-99 Length: 978
 Score: 944.50 Matches: 186
 Percent Similarity: 83.0% Conservative: 39
 Best Local Similarity: 68.6% Mismatches: 43
 Query Match: 45.9% Indels: 3
 DB: 7 Gaps: 1

US-10-618-570-2 (1-400) x CN643776 (1-978)

Qy 1 MetAlaGlnTrpAspAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
 Db 169 ATGGAGCAGTGGGATCGCTTTCGTGATCAACAGAGGACGTTGATGCTGTCTGAATCT 228

Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
 Db 229 GTGAAATTTGAGCTCGCTCAATGACAGCTTTGCTTCTCTCTAATCTTAAACACAGCCCT 288

Qy 41 ThrLeuGlnGluArgMetLysSerTyrlsThrAlaLeuLeuThrLeuTyrlsLeuVal 60
 Db 289 CCCCTTCAAGAGAACTGAAGTCTTCAAGAGTCTCACTGATTCCTTACCTCTCTGCTG 348

Qy 61 PheValValLeuValProIleGlyLeuValAlaAlaGlnLeuLeuLysTrpGluThr 80
 Db 349 TTTCAGTTCATCTCCCTCTCATTTGGAATAGTGGCAGCTCACTCTGAAATGGGAACG 408

Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99

```
Db 409 AAGAAATTCCTCCATTGGTTCTACTAATGCAGCAGCATATAACTCAAGAGTCTCACAGGAAA 468
Qy 100 GlyAenGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAen 119
Db 469 GGAATGACATGAGCGGACGAGATTTTCAAGAAAGTCTTTATGGAAACACATGAGCAAC. 528
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAenGluAlaAenLeuLeuAspAlaLysAen 139
Db 529 ATGGAGAAGAGATCCAGCACATTTCCAGCATGAGCAATCTCATAGATGAGAGCAT 588
Qy 140 PheGlnAenPheSerIleThrAspGlnArgPheAenAspValLeuPheGlnLeuAen 159
Db 589 TTCCAAAACCTTCAGCATGACAACTGATCAAGATTTAATGACATTTCTTCTCCAGTAAGT 648
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAenIlelleGlyAspIleSerLysSerLeu 179
Db 649 ACCTTGTTTCTCAGTCCAGGACATGGGAATACAATAGATGAATCTCCAAAGTCTCTTA 708
Qy 180 ValGlyLeuAenThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAenGlyArg 199
Db 709 ATAAGTTTGATACCACTGCTGTTGATTTGGAGCTCAACATAGAAAGCTGAATGGCAAA 768
Qy 200 ValGlnGluAenAlaPheLysGlnGlnGluMetArgLysLeuGluGluArg-IleTy 219
Db 769 ATCCAAGAGAAAACCTTCAACACACAGAGGAAATCAGTAAATTAGAGGACATTGTAA 828
Qy 219 rAsnAlaSer-AlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnI 239
Db 829 TAAATGTTATCCAGCAGAAATTTATGCTATGAAAGAGAAACAAGTGCATTTGGAAACAGAAA 888
Qy 239 leLysGlyGluMetLysLeuAenAenIleThrAenAspLeuArgLeuLysAspTrpG 259
Db 889 TAAAGGGAATTAAGTACTGATATACATCACCATGATCTCAGACTGAAGACTGGG 948
Qy 259 luHisSerGlnThrLeuLysAenIleThr 268
Db 949 AACATTTCTCAAACTTGGAGAAATATCC 977

RESULT 7
BX496563
LOCUS
DEFINITION
  DKFZp779D0328 r1 779 (synonym: hnccl) Homo sapiens cDNA clone
  DKFZp779D0328 5', mRNA sequence.
ACCESSION
  BX496563
VERSION
  BX496563.1 GI:32012303
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Organism
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Homnidae; Homo.
REFERENCE
  1 (bases 1 to 843)
  Ottenwaelder B., Obermaier, B., Deutschenbaur, S., Mewes, H.W.,
  Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.
  EST (Ottenwaelder, B., Obermaier, B., Deutschenbaur, S., Mewes, H.W.,
  et al.)
  Unpublished (2003)
  Contact: MIPS
  MIPS
  Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
  This is the 5' sequence of the clone insert
  Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
  Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
  sequenced by MedGenomix (Martinsried/Germany) within the CDNA
  sequencing consortium of the German Genome Project. No s1 sequence
  available.
  This clone (DKFZp779D0328) is available at the RZPD in Berlin.
  Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
  Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
  Location/Qualifiers
    1..843
    /organism="Homo sapiens"
```

```
ORIGIN
Alignment Scores:
Pred. No.: 1.1e-92 Length: 843
Score: 885.00 Matches: 182
Percent Similarity: 84.0% Conservative: 39
Best Local Similarity: 69.2% Mismatches: 40
Query Match: 43.0% Indels: 5
DB: 5 Gaps: 1

US-10-618-570-2 (1-400) x BX496563 (1-843)
Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
Db 58 ATGAGCAGTGGGATCACTTTTCAACATCAACAGAGGACACTGATAGTCTCCGAATCT 117
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAenGlyPro 40
Db 118 GTGAATTTGATGCTCGCTCAATGACAGCTTTGCTTCTCCGAATCTTAAACACAGCCT 177
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60
Db 178 TCCCTTCAAGAGAACTGAAGTCTTCAAGCTGCACCTGATTTGCCCTTTTACTCTCTCGTG 237
Qy 61 PheValValLeuValProIlelleGlylleValAlaAlaGlnLeuLeuTyrTrpGluThr 80
Db 238 TTTGCACTTCTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTGAAGTGGGAACG 297
Qy 81 LysAenCysThrValGlySerValAenAla---AspIleSerProSerProGluGlyLys 99
Db 298 AAGATTTGCTCAGTTAGTTCACTAATGCAATGATATATACTCAAGTCTCACGGGAAA 357
Qy 100 GlyAenGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAen 119
Db 358 GGAATGACACGGAAGAGAAATGAGATTTTCGAGAAGCTTTTATGGAAACACATGAGCAAC 417
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAenGluAlaAenLeuLeuAspAlaLysAen 139
Db 418 ATGGAGAAGAGAAATCCAGCATATTTTAGACATGGAAGCCCACTCATGGACACAGAGCAT 477
Qy 140 PheGlnAenPheSerIleThrAspGlnArgPheAenAspValLeuPheGlnLeuAen 159
Db 478 TTTCCAAAATTTTCCAGCATGACAACTGATCAAGATTTAATGACATTTCTTCTGCAGTAAGT 537
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAenIlelleGlyAspIleSerLysSerLeu 179
Db 538 ACCTGTTTCTCAGTCCAGGACATGGAATGCAATAGATGAAATCTCCAACTCTCTTA 597
Qy 180 ValGlyLeuAenThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAenGlyArg 199
Db 598 ATAAAGTTTGAATACCACTTGTGATTTGCACTCAACATAGAAATCTGAATGGGAAA 657
Qy 200 ValGlnGluAenAlaPheLysGlnGlnGluMetArgLysLeu-GluGluArgIleTy 219
Db 658 ATCCAGAGATACCTTCAACACACAGAGAAATCAGTAAATTAGAGGAGCGGTGTTA 717
Qy 219 rAsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGlnGluI 239
Db 718 CAATGTATCAGCAGAAATTTATGGCTATGAAAGAGAAACAAGTGCATTTGGAACAGGAAT 777
Qy 239 eLysGlyGluMetLysLeuLeuAenAenIleThrAenAspLeuArgLeuLysAspTrpG 259
Db 778 -AAGGAGAGAGGAAAGTACTGAATAC-ATCNCNTAATGATCTCAGACTGGAAGA-TGGGA 834
Qy 259 uHisSer 261
```

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp779D0328"
/tissue_type="liver"
/dev_stage="fetal"
/lab_host="DH10B"
/clone_lib="779 (synonym: hnccl)"
/notes="Vector: pSport1_Sfi, Site_1: SfiIA, Site_2: SfiIB"
```

```

Db      835 ACATTCT 841
|||||
RESULT 8
AK028480
LOCUS
DEFINITION
Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched
library, clone:4631424C05 product:macrophage scavenger receptor 1,
full insert sequence.
ACCESSION
AK028480.1 GI:26324437
VERSION
AK028480
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6
(bases 1 to 3636)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hirao, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saio, R., Saichou, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers
1. 3636
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:4631424C05"
/db_xref="taxon:10090"
/clone="4631424C05"
/tissue_type="skin"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
53. 1117
/note="unnamed protein product; macrophage scavenger
receptor 1 (MGD) MGI:98257, GB|NM_031195, evidence: BLASTN,
99% match=1461)
putative"
/codon_start=1
/protein_id="BAC25973.1"
/db_xref="GI:26324438"
/translacion="MTKEMTENVNQLCPHEQEDADCSSESVKFDRSMTASLPHSTKNG
PSLOELKSPKALIALYLVFVLPVGVITVQALLNWMKNCVLSINTSDTSQGP
MEKTSKVENRFTLIIMHMKMBERIESNSKADLIDTERFQNFQNSVQDRLNDIL
LQNLNLSVSSVQEHGNSDASKLSQNLNMLDQLTETLNRVSTAKQEDISK
LEERYKVSAREVQKEQAEVQEVRLNNITNDLRKQWHSQTLKNTFFIQ
GPPQPEKGGDRGLTGOTGPPGPGIRGIPGVKDRGQIGFPGGRNGPCAPKPKRSG
SPGPKGKGKSGSVGSSRSV"
ORIGIN
Alignment Scores:
Pred. No.: 2.93e-85 Length: 3636
Score: 830.00 Matches: 162
Percent Similarity: 78.6% Conservative: 47
Best Local Similarity: 60.9% Mismatches: 55
Query Match: 40.3% Indels: 2
DB: 4 Gaps: 2
US-10-618-570-2 (1-400) x AK028480 (1-3636)
Qy 8 ProaspGlnGlnGluAapThrAspSerCysThrGluSerVallysPheAspAlaArgSer 27
Db 89 CCTCATGAACAAGAGGATGCTGACTGCAGTTCAGAAATTCGACGACGCTTCA 148
Qy 28 ValThrAlaLeuLeuProHisProlysaAsnGlyProThrLeuGlnGluArgMetLys 47
Db 149 ATGACTGCATCCCTTCCTCAGCAGCACTAAATGGCCCTCCCTTCAGAGAGTTGAAG 208
Qy 48 SerTyrIysThrAlaLeuLeuThrLeuTyrLeuLeuValPheValLeuValProLe 67
Db 209 TCCTTCAAGGCTGCCCTCATTTGCTCTCTACCTCTCTGTTTGGCAGTACTAATACCTGT 268
Qy 68 IledlylleValAlaAlaGlnLeuLeuLysTrpGluThrLysAsnCythrValGlySer 87
Db 269 GTTGGAAATAGTAACAGCTCAGCTTTTGAATTTGGGAAATGAAGAACTGCTTGTAGTTT 328
Qy 88 ValaAsn--AlaAapilleSerProSerProGluGlyLysGlyAsnGlySerGluAaspGlu 106
Db 329 CTTAAACAGATGACACATCTCAAGGTCCTATGAAAGAAAGAAATACACAGTAAGTGGAA 388
Qy 107 MetArgPheArgGluAlaValMetGluArgMetSerAsnMetGluSerArgileGlnTyr 126
Db 389 ATGAGATT---ACAATTATCATGCAACACATGAAGCAGCATGGAGGAGAGAAATCGAAGC 445
Qy 127 LeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPheGlnAsnPheserIleThr 146
Db 446 ATTTCAAACTCAAAAGCGGACCTTTATACACACGGAACGCTTCCAGAAATTTCCAGCATGCA 505

```

Qy 147 ThrAspGlnArgPheAsnAspValLeuPheGlnLeuAenSerLeuLeuSerSerIleGln 166
 Db 506 ACTGACCAAGACATTAATGATATTTCTTCGAGTTAAATTTCTTGATTTCTGAGTCAGTCAG 565

Qy 167 GluHisGluAenIleGlyAspIleSerLeuValGlyLeuAenThrThrVal 186
 Db 566 GAACATGGGAATTCACCTGGATGCATCTCCAAAGTCCTTGCAGAGTCTGAATATGACACTG 625

Qy 187 LeuAspLeuGlnPheSerIleGlnThrLeuAenGlyArgValGlnGlnAenAlaPheLys 206
 Db 626 CTTGATGTTCAACTCATACAGAACTGAATGTGCAGAGTCCGTGATCTACAGCAAG 685

Qy 207 GlnGlnGluGluMetArgLysLeuGluGluArgIleThrAenAlaSerAlaGluIleLys 226
 Db 686 CAACAGGAGGACATCAGTAAATTTGGAGAACTGTGTACAAAGTATCATCAGAGAGTCCAG 745

Qy 227 SerLeuAspGluLysGlnValTyrLeuGluGlnGluIleLysGlyGluMetLysLeuLeu 246
 Db 746 TCTGTGAAGAGAACACAGCGCAGCTGGAAACAGGAGTAAACAGGAGTGCAGATATTG 805

Qy 247 AenAenIleThrAsnAspLeuArgLeuLysAspTrpGluHisSerGlnThrLeuLysAen 266
 Db 806 AACACATCACCACGACTCAGACTGAAGAGTGGGAACACTCAGACACTGAAAAAC 865

Qy 267 IleThrLeuLeuGlnGly 272
 Db 866 ATCACCTTCATTCAAGGG 883

RESULT 9
 AY413210
 LOCUS Mus musculus MSRI gene, VIRUAL TRANSCRIPT, partial sequence,
 DEFINITION genomic survey sequence.
 ACCESSION AY413210
 VERSION AY413210.1 GI:39769175
 KEYWORDS GSS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.
 1 (bases 1 to 1030)
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302

REFERENCE
 AUTHORS 2 (bases 1 to 1030)
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.

TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.

FEATURES
 source
 1..1030
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 <1..>1030
 /gene="MSRI"
 /locus_tag="HMC4814"

ORIGIN
 Alignment Scores: 2.84e-85 Length: 1030
 Pred. No.:

Score: 823.00 Matches: 160
 Percent Similarity: 78.6% Conservative: 49
 Best Local Similarity: 60.2% Mismatches: 55
 Query Match: 40.0% Indels: 2
 DB: 10 Gaps: 2

US-10-618-570-2 (1-400) x AY413210 (1-1030)

Qy 8 ProAspGlnGlnGluAspThrAspSerCysThrGluSerValLysPheAspAlaArgSer 27
 Db 22 CCTCATGACGAGAGGATGCTGACTGCAGTTCAAGATCCGTAATTTTGACGCGACGTTCA 81

Qy 28 ValThrAlaLeuProHisProLysAenGlyProThrLeuGlnGlnArgMetLys 47
 Db 82 ATGACAGATCCCTTCTCTCACAGCACTTAAATATGCCCCCTCCGTTCAGGAGAGTTGAG 141

Qy 48 SerTyrLysThrAlaLeuIleThrLeuTyrLeuIleValPheValValLeuProfile 67
 Db 142 TCCTTCAAGGTCGCCCTCATTTGCTCTCTACCTCTGTGTTTGCAGTACTAATACCTGTT 201

Qy 68 IleGlyIleValAlaAlaGlnLeuLeuLysTrpGluThrLysAenCysThrValGlySer 87
 Db 202 GTTGGAAATAGTCAGCTGAGCTTTTGAATTTGGGAAATGAAGAACTGCTTAGTTGTTCA 261

Qy 88 ValAen--AlaAspIleSerProSerProGluGlyLysGlyAenGlySerGluAspGlu 106
 Db 262 CTTAAACAAGTGACACATCTCAAGTCTCTATGAGAAAGAAATACCAAGTAAAGTGAA 321

Qy 107 MetArgPheArgGluAlaValMetGluArgMetSerAenMetGluSerArgIleGlnTyr 126
 Db 322 ATGAGATTT--ACAATTTATCATGGAACACATGAAGGACATGGAGGAGAGAATCGAAAGC 378

Qy 127 LeuSerAspAenGluAlaAenLeuLeuAspAlaLysAenPheGlnAenPheSerIleThr 146
 Db 379 ATTTCAAACTCAAAGCGCGACCTTATAGACACGGAACGCTTCCAGAAATTTCAGCATGGCA 438

Qy 147 ThrAspGlnArgPheAenAspValLeuPheGlnLeuAenSerLeuLeuSerSerIleGln 166
 Db 439 ACTGACCAAGACTTAATGATATTTCTTCTGCAGTTAAATTTCTTGATTTCTGAGTCAG 498

Qy 167 GluHisGluAenIleGlyAspIleSerLysSerLeuValGlyLeuAenThrThrVal 186
 Db 499 GAACATGGGAATTCACCTGGATGCAATCTCCAAGTCTTGCAGAGTCTGAATATGACACTG 558

Qy 187 LeuAspLeuGlnPheSerIleGluThrLeuAenGlyArgValGlnGlnAenAlaPheLys 206
 Db 559 CTTGATGTTCAACTCCATACAGAAACACTGAATGTGCAGAGTCCGTGAACTTACAGCAAG 618

Qy 207 GlnGlnGluGluMetArgLysLeuGluGluArgIleTyrAenAlaSerAlaGluIleLys 226
 Db 619 CAACAGGAGGACATCAGTAAATTTGAGGAACCTGTGTACAAAGTATCATCAGGAGAGTCCAG 678

Qy 227 SerLeuAspGluLysGlnValTyrLeuGluGlnGluIleLysGlyGluMetLysLeuLeu 246
 Db 679 TCTGTGAAGAGAAACAAGCGCACGTCGAAACAGGAAGTAAACAGGAAGTGCAGAGTATTG 738

Qy 247 AenAenIleThrAenAspLeuArgLeuLysAspTrpGluHisSerGlnThrLeuLysAen 266
 Db 739 AACAACTATCAACAGGACCTCAGACTGAAGACTGGGAACACTCAGACACTGAAAAAC 798

Qy 267 IleThrLeuLeuGlnGly 272
 Db 799 ATCACCTTCATTCAAGGG 816

RESULT 10
 AK089178
 LOCUS Mus musculus NOD-derived CD11c +ve dendritic cells cDNA, RIKEN
 DEFINITION full-length enriched library, clone: F630005M03 product: macrophage
 scavenger receptor 1, full insert sequence.
 ACCESSION AK089178
 VERSION AK089178.1 GI:26354301
 KEYWORDS HTC; CAP trapper.

Qy 167 GluHisGluAenIleGlyAspIleSerLysSerLeuValGlyLeuAsnThrThrVal 186
 Db 566 GAACATGGGAATTCACCTGGATGCATCTCCCAAGTCCTTGCAGAGTCTGAATATGACACTG 625
 Qy 187 LeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgValGlnGluAsnAlaPheLys 206
 Db 626 CTTGATGTTCACTCCATACAGAACTGTCATGTCAGAGTCCGTGAATCTACAGCAAG 685
 Qy 207 GlnGlnGluGluMetArgLysLeuGluGluArgLysLeuAsnAlaSerAlaGluLys 226
 Db 686 CAACAGGAGGACATCAGTAAATTCAGGAGAACTGTGTACAAAGTATCATCAGCAAGTCCAG 745
 Qy 227 SerLeuAspGluLysGlnValThrLeuGlnGluLysGlyGluMetLysLeuLeu 246
 Db 746 TCTGTGAAGAGAAACAAGCCGACCTGGAAAGGAGTAAACAGGAAGTGGAGTATTG 805
 Qy 247 AsnAenIleThrAsnAspLeuArgLysLeuAspTrpGluHisSerGlnThrLeuLysAsn 266
 Db 806 AACACATCACCAAGCACTCAGACTGAAGACTGGGAACACTCACAGCACTGAAAAAC 865
 Qy 267 IleThrLeuLeuGlnGly 272
 Db 866 ATCACCTTCATTCAAGGG 883

RESULT 11
 BW956649
 LOCUS
 DEFINITION
 BW956649 full-length enriched swine cDNA library, adult alveolar
 macrophage Sus scrofa cDNA clone AMP010022D07 5', mRNA sequence.
 ACCESSION
 BW956649
 VERSION
 BW956649.1 GI:71974090
 KEYWORDS
 EST.
 SOURCE
 Sus scrofa (pig)
 ORGANISM
 Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
 Sus.
 1 (bases 1 to 677)
 Uenishi,H., Eguchi,T., Suzuki,K., Sawazaki,T., Toki,D., Shinkai,H.,
 Okumura,N., Hamasima,N. and Awata,T.
 TITLE
 PEDE (Pig EST Data Explorer): construction of a database for ESTs
 derived from porcine full-length cDNA libraries
 JOURNAL
 Nucleic Acids Res. 32 (1), D484-D488 (2004)
 PUBMED
 14681463
 COMMENT
 Contact: Hirohide Uenishi
 Animal Genome Laboratory, Genome Research Department
 National Institute of Agrobiological Sciences
 2 Ikenodai, Tsukuba, Ibaraki 305-8602, Japan
 Tel: +81-29-838-8627
 Fax: +81-29-838-8627
 Email: huenishi@affrc.go.jp
 EST project with full-length enriched cDNA libraries carried out in
 Animal Genome Research Program (Japan) by National Institute of
 Agrobiological Sciences and STAFF-Institute
 Single pass sequencing of clones derived from oligo-capped cDNA
 library

Vector sequences were eliminated by RepeatMasker version 2002/07/13
 and crossmatch version 0.990319
 Low quality bases were trimmed based on the quality values.
 FEATURES
 Location/Qualifiers
 1..677
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /clone="AMP010022D07"
 /tissue_type="alveolar macrophage"
 /dev_stage="adult"
 /clone_lib="full-length enriched swine cDNA library, adult
 alveolar macrophage"

ORIGIN

Alignment Scores: 7.46e-82 Length: 677
 Pred. No.:

Score: 791.50 Matches: 152
 Percent Similarity: 84.4% Conservative: 26
 Best Local Similarity: 72.0% Mismatches: 26
 Query Match: 38.4% Indels: 7
 DB: 5 Gaps: 1
 US-10-618-570-2 (1-400) x BW956649 (1-677)
 Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
 Db 59 ATGGAACAGATGGCCATACCTTTCTCGATCAACCCGAGGACACTGACAGTTGTATGGAATCT 118
 Qy 21 VallysPheAspAlaArgSerValThrAlaLeuLeuProProHisProLysAsnGlyPro 40
 Db 119 GTCAAGTTCATGCTCGCTCAATGACAGCTTGTCTCTCAATCTCTAATAATGGCCCA 178
 Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLysLeuThrLeuLysVal 60
 Db 179 ACGCTTCAAGAGAAAATGAAGTCTTTCAAAAGTCGACTGATTGGCCCTTTATCTCTCTGTG 238
 Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysThrGluThr 80
 Db 239 TTTGTGCTGCTGGTGCCTATCATCGAGTACTCGACGCTCATCTCTGAAATGGGAATG 298
 Qy 81 LysAsnCysThrValGlySerValAsnAlaAspIleSerProSerProGluGlyLysGly 100
 Db 299 AAGAAATTGTCAGTTGGCTCAATTAGTGCA-----GGAAAGGA 337
 Qy 101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120
 Db 338 AATGACAGTGAATGACATGAGATTTTCGTGACGCTGTACAGAAACATATGAGCAACATG 397
 Qy 121 GluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPhe 140
 Db 398 GAGAAGAGAAATTTAGTATCTTTTCAGATAATGAAGCCCAATCTCGTAGAATCCGAGCATTTTC 457
 Qy 141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
 Db 458 CAAAATTTCTACATGCATGCTCAAGATTTTAATGATGTCCTTTCCAGCTTAATAAC 517
 Qy 161 LeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeuVal 180
 Db 518 TTGCTTTCTTCATTACGGAACATGGAAATGTAATAAATGAAATCTCTAAGTCATTATA 577
 Qy 181 GlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal 200
 Db 578 AGTCTGAACACCACTTGTGATTTGAGCTCAATTCGAGTCAATTCGAAACACTGAAAGGCAAGTC 637
 Qy 201 GlnGluAsnAlaPheLysGlnGluGluMet 211
 Db 638 CAAGAGAAATCAATTTAAACAACAGAGAGATG 670

RESULT 12

DR004903

LOCUS

DEFINITION

DR004903

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

DR004903 768 bp mRNA linear EST 17-MAY-2005
 TCl09459 Human placenta, large insert, PCMV expression library Homo
 sapiens cDNA clone TCl09459 5' similar to Homo sapiens macrophage
 scavenger receptor 1 (MSR1), transcript variant SR-A11, mRNA
 sequence.
 DR004903
 DR004903.1 GI:66264776
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 768)
 Birkett,C., Cho,J., Gau,Y., Hamer,R., Kelly,S., Kovacs,K., Liu,L.,
 Liu,X., Porter,J., Sachs,A., Shu,Y., Sun,Z., Wong,J., Wu,M.,
 Zhang,X., Jay,G. and He,W.
 High-throughput cloning of full-length human cDNAs directly from
 cDNA libraries optimized for large and rare transcripts

JOURNAL
COMMENT

Unpublished (2005)

Contact: Kovacs, KF
High throughput cDNA Cloning
Origene Technologies, Inc. (www.origene.com)
6 Taft Court, Suite 100, Rockville, MD 20850, USA
Tel: 301 340 3188
Fax: 301 340 8606
Email: cDNA@origene.com

This EST submission is part of an on-going human full-length
cloning project at Origene Technologies, Inc.

Please contact Origene for access.

Origene Technologies, Inc.

6 Taft Ct. Suite 100

Rockville, MD 20850

Tel: (301) 340-3188

http://www.origene.com

Seq primer: pCMV6 5prime forward vector primer, OriGene
Technologies Inc.

FEATURES
source

Location/Qualifiers
1..768
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TC109459"
/tissue_type="Placenta"
/clone_lib="Human placenta, large insert, pCMV expression
library"
/note="Vector: pCMV6-XL4; Site 1: EcoRI; Site 2: XhoI/Sall
compatible end ligation; Oligo-dT primed reverse
transcription optimized for large and GC rich mRNA
transcripts, cDNA size selection, optimized ligation for
large inserts into mammalian expression vector, random
clones selected for end sequence verification of
full-length genes"

ORIGIN

Alignment Scores:

Pred. No.: 4,54e-79 Length: 768
Score: 768.50 Matches: 146
Percent Similarity: 84.4% Conservative: 32
Best Local Similarity: 69.2% Mismatches: 32
Query Match: 37.3% Indels: 1
DB: 8 Gaps: 1

US-10-618-570-2 (1-400) x DR004903 (1-768)

Qy 1 MetAlaGlnTrpAspAspPheProAspGlnGluAspThrAspSerCysThrGluSer 20
Db 123 ATGGAGCGATGGGATCACTTTTCACAACTCAACAGGAGGACCTGATAGCTGCTCGAATCT 182
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 183 GTGAATTTGATGCTCGCTCAATGACAGCTTTGCTTCTCCGATCTTAAACACAGCCT 242
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
Db 243 TCCCTTCAAGAGAACTCAAGTCTTCAAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 302
Qy 61 PheValValLeuValProIleGlyLeuValAlaGlnLeuLeuTyrTrpGluThr 80
Db 303 TTTGAGTTCATCCCTCTATTGGAATGTCGAGCTCACTCACTCACTCACTCACTCACTCACT 362
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 363 AAGATTGCTAGTTAGTTCACTAATGCAATGATATGATATCAAGTCTCACGGGAAA 422
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 423 GGAATGACAGCGAAGAGAAATGAGATTTCAAGAAGTCTTTATGGAACACATGAGCAAC 482
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 483 ATGGAGAGAGAAATCCAGCATATTTTAGCATGGAAGCCAACTCTCATGGACACAGCAT 542

Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 543 TTCCAAAATTTTCAGCATGACAACTGATCAAAAGATTTAATGACATTTCTTCAGCTAAGT 602
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
Db 603 ACCTTGTTTCTCAGTCCAGGACATGGGAATGCAATAGATGAATCTCCAAGTCCTTA 662
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 663 ATAAGTTTGTATACCACTTCCTTGTATTCAGCTCAACATAGAAAATCTGATGGCAA 722
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluGlu 210
Db 723 ATCAAGAGATACCTTCAAAACACAGAGAAA 755

RESULT 13

CB469554 LOCUS CB469554 626 bp mRNA linear EST 26-MAR-2003
DEFINITION sn09.A05.f sn Sus scrofa cDNA 5', mRNA sequence.
ACCESSION CB469554
VERSION CB469554.1 GI:29275939
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
Sus.
REFERENCE 1 (bases 1 to 626)
AUTHORS Neilan,J.G., Kutish,G.F., Lu,Z., Zsak,A. and Rock,D.L.
TITLE Sequence analysis of African swine fever virus infected and
non-infected porcine macrophage cDNA libraries
JOURNAL Unpublished (2003)
COMMENT Contact: Neilan JG
Plum Island Animal Disease Center
US Department of Agriculture, Agricultural Research Service
PO Box 848, Greenport, NY 11944-848, USA
Tel: 631 323 3133
Fax: 631 323 3044
Email: jneilan@piadc.ars.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim.alt option. Vector identified by
cross_match v0.990329 and lucy vl.17p.
Seq primer: M13 Forward.

FEATURES

source

Location/Qualifiers
1..626
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="lymphoid"
/cell_type="macrophage"
/lab_hosts="DH10B"
/clone_lib="sn"

/note="Vector: pSPORT1; Site 1: NotI; Site 2: Sall;
Library made from pools of polyA selected RNA, isolated at
different times post-infection (0 to 16 hrs) from African
swine fever virus (isolate Pretoriuskop/96/4) infected
swine macrophages. Macrophages were derived from
peripheral blood mononuclear cells cultured for 48 hrs on
plastic in the presence of 30% L929 supernatant."

ORIGIN

Alignment Scores:
Pred. No.: 1.9e-75 Length: 626
Score: 736.50 Matches: 141
Percent Similarity: 84.7% Conservative: 25
Best Local Similarity: 71.9% Mismatches: 23
Query Match: 35.8% Indels: 7
DB: 6 Gaps: 1

US-10-618-570-2 (1-400) x CB469554 (1-626)

QY 200 ValGlnGlu 202
 Db 643 ATCCAAGAG 651
 RESULT 15
 CN643736
 LOCUS
 DEFINITION
 ILLUMIGEN_MCO_8498 Katze_MMR Macaca mulatta cDNA clone IBIUW:3182
 5' similar to Bases 116 to 1087 highly similar to human MSR1
 (Hs.436887), mRNA sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Macaca mulatta (rhesus monkey)
 Macaca mulatta
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Cercopithecoidea; Cercopithecinae; Macaca.
 REFERENCE
 AUTHORS
 Magness, C.L., Fellin, P.C., Thomas, M.J., Korth, M.J., Agy, M.B.,
 Prohl, S.C., Fitzgibbon, M., Scherer, C.A., Miner, D.G., Katze, M.G. and
 Iadonato, S.P.
 TITLE
 Analysis of the Macaca mulatta transcriptome and the sequence
 divergence between Macaca and human
 JOURNAL
 PUBMED
 Genome Biol. 6 (7), R60 (2005)
 COMMENT
 15998449
 Contact: C. Magness
 Illumigen Biosciences Inc.
 2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
 Tel: 2063780400
 Fax: 2063780408
 Email: cmagness@illumigen.com
 Sequenced on 2004.01.15, 487 Q20 bases.
 PCR Primers
 FORWARD: CCCTCACTAAAGGGAACAAA
 BACKWARD: CACTATAGGCGGAATTGGTA
 Insert Length: 1093 Std Error: 0.00
 Plate: CL000089 row: C column: 03
 Seq primer: CCCTCACTAAAGGGAACAAA
 POLYA=Yes.

Location/Qualifiers
 1..1093

/organism="Macaca mulatta"
 /mol_type="mRNA"
 /strain="Indian"
 /db_xref="taxon:9544"
 /clone="IBIUW:3182"
 /sex="female"
 /dev_stages="adult"
 /lab_host="E. coli SOLR"
 /clone_lib="Katze MMR"
 /note="Organ: brain; Vector: Uni-ZAP XR; Site 1: Rcor 1;
 Site 2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
 kit (Catalog #200400) and ZAP-cDNA Gigapack III Gold
 Cloning Kit (Catalog #200450)"

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.: 4.4e-73 Length: 1093
 Score: 719.50 Matches: 162
 Percent Similarity: 73.6% Conservative: 44
 Best Local Similarity: 57.9% Mismatches: 64
 Query Match: 34.9% Indels: 10
 DB: 7 Gaps: 3

US-10-618-570-2 (1-400) x CN643736 (1-1093)

QY 1 MetAlaGlnTrpAspProAspGlnGluAspThrAspSerCysThrGluSer 20
 Db 237 ATGGAGCAGTGGATCGCTTCGTGATCAACAGAGGACGTTGATAGCTGCTGTAATCT 296
 QY 21 ValLysPheAspAlaArgSerValThrAlaLeu-LeuProHisProLysAsnGlyPr 40

Db 297 GTGAAATTTGACGCTCGCTCAATGACAGCTTTCCTCTATCTTAAACAGCCC 356
 QY 40 oThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuLeuThrLeuTyrLeuIleVa 60
 Db 357 TCCCTTTCAAGAGAAACTGAAGTCTCTCAAGAGTGCACCTGATGTCCTTTACCTCTTGT 416
 QY 60 lPheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTrpGluTh 80
 Db 417 GTTTGCGATTTCTCATCCCTCTCATTTGGAATAGTGCAGCTCAGCTTCTGAAGCGGAAAC 476
 QY 80 rLysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLy 99
 Db 477 GAAGAAATTCCTCATTTGTTCTTACTAATGCAGACGATATAACTCAAGTCTCACAGGAAA 536
 QY 99 sGlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAs 119
 Db 537 AGGAATACACAGTGAAGCGGAACGAGATTTCAAGAGTCTTTATGGAAACACATGAGCAA 596
 QY 119 nMetGluSerArgIleGlnTyrLeuSer-AspAsnGluAlaAsnLeuLeuAspAlaLysA 139
 Db 597 CATGGAGAAGAGAAATCCAGCACATTTATACATGTAGGCCAATCTCATAGATGCAGAGC 656
 QY 139 snPheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuA 159
 Db 657 ATTTCCAAAACTTCAGCATGACAACTGATCAAGATTTAATGACATCTCTCCAGCTAA 716
 QY 159 snSerLeu-LeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSe 178
 Db 717 GTACTTGTGTTTTCCTCAGTCCAGCGNATGGGATACATAGATGAATCTCCAGTGC 776
 QY 178 rLeuValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGl 198
 Db 777 CTTAATAAGCTTGAATACCACTTTGCTTGTATTGAGCTCAACATAAAAAAGCTGAATGG 836
 QY 198 YArgValGlnGluAsnAlaPheLysGlnGlnGluMetArgLys--LeuGluGluArg 217
 Db 837 CAAAATCCAGAAAAAACCTTCAAAACCCAGATGTATATCCAGAAAAATTTAGGGGAGCAT 896
 QY 218 lIleTyrAsnAla---SerAlaGluIleLysSerLeuAspGluLysGln---ValTyrLeu 235
 Db 897 GTTTATATGATTCCTCGCTTAACCTTATGTTTGAAGAGAACAGCAGGCTCTTTTGTGA 956
 QY 236 GluGlnGluIleLysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeu 255
 Db 957 ACAGGAAAAAAGAGATAAGGGGAAAGACCTTAAAAAAATTTCCCAATGATCTCAAAATG 1016
 QY 256 -LysAspTrpGluHisSerGlnThrLeuLysAsnIleThrLeuLeu 270
 Db 1017 GAAAACCTGGGAAATTTCCAGAAATCTTGGAAAAAAATTTCTTTTA 1062

Search completed: February 27, 2006, 23:22:58
 Job time : 5155 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 27, 2006, 08:46:22 ; Search time 646 Seconds
(without alignments)
4126.745 Million cell updates/sec

Title: US-10-618-570-2

Perfect score: 2059

Sequence: 1 MAQWDFPDQEDTDSCTES.....DWKATRVGINFTRLRTQKE 400

Scoring table:

Bgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/absq/ABSSWB pool/US10618570/runat 27022006 063853 13038/app query fasta_1
-DB=N Geneseq -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=spct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs05p
-USER=US10618570 @CGN 1 1 1096 @runat 27022006 063853 13038 -NCPU=6 -ICPU=3
-NO.MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 21:*

1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2059	100.0	5177	2	Az09997 Bovine sc
2	1381.5	67.1	1588	2	Az04926 Sequence
3	1381.5	67.1	1588	2	Az028540 DNA for b
4	1381.5	67.1	1807	2	Az04925 Sequence

5	1042.5	50.6	1077	10	ADB89012	Human DNA
6	1042.5	50.6	1347	2	AAQ21550	Gene enco
7	1042.5	50.6	1347	2	AAQ47731	Human sca
8	1042.5	50.6	1347	4	AAF29165	Human DNA
9	1042.5	50.6	1367	2	AAT43260	Type II m
10	1042.5	50.6	1367	10	ACA56662	Human sig
11	1042.5	50.6	1367	12	ADI56458	Human pol
12	1042.5	50.6	2028	2	AAQ21549	Gene enco
13	1042.5	50.6	2028	2	AAQ47730	Human sca
14	1042.5	50.6	2028	2	AAT43259	Type I ma
15	1042.5	50.6	2028	4	AAF29164	Human DNA
16	1042.5	50.6	2028	12	ADI62122	Human inac
17	1042.5	50.6	2877	13	ADQ38907	Human SNP
18	1042.5	50.6	3336	13	ADQ38905	Human SNP
19	1042.5	50.6	3525	13	ADQ38904	Human SNP
20	1042.5	50.6	3704	6	AAS94868	Human DNA
21	1042.5	50.6	3719	13	ADQ38906	Human SNP
22	1042.5	50.6	3791	13	ADQ38903	Human SNP
23	1042.5	50.6	3898	13	ADP24728	PRO polyp
24	1039.5	50.5	2037	2	AAT68796	Macrophag
25	1032.5	50.1	2037	2	AAQ28539	DNA for h
26	994.5	48.3	1330	12	ADG39681	Human cDN
27	994.5	48.3	1330	12	ADG39730	Human cDN
28	688	33.4	459	12	ADP86098	Barley al
29	688	33.4	484	2	AAT68502	Barley al
30	680.5	33.1	1522	12	ADM97492	CD1d-1gG-
31	678.5	33.0	1746	14	AEA28083	Circularl
32	678	32.9	387	14	ADV78150	Chicken a
33	678	32.9	459	2	AAZ80197	Chicken a
34	678	32.9	480	14	ADV78151	Barley al
35	678	32.9	486	3	AAZ49865	Potato pr
36	678	32.9	576	3	AAZ49855	pGEMav av
37	678	32.9	604	2	AAT73194	Wild-type
38	678	32.9	604	12	ADP26979	Chicken a
39	678	32.9	823	12	ADM97488	CD1d-1gG-
40	678	32.9	871	12	ADM97490	CD1d-1gG-
41	677.5	32.9	873	14	AEA28087	Circularl
42	670	32.5	540	14	ADX98285	Chicken a
43	670	32.5	540	14	ADX98286	Chicken a
44	647	31.4	401	12	ADO85120	DNA seque
45	612	29.7	433	12	ADO85126	DNA seque

ALIGNMENTS

RESULT 1

AAZ09997

ID AAZ09997 standard; cDNA; 5177 BP.

XX

AC AAZ09997;

XX

DT 03-DEC-1999 (first entry)

XX

DE Bovine scavenger receptor class A (ScR)/avidin fusion protein cDNA.

XX

KW Scavenger receptor class A; ScR; avidin; fusion protein; bovine; ECD;
KW membrane-spanning domain; extracellular domain; biotin-binding activity;
KW endocytosis; ss.

XX

OS Synthetic.

OS Bos taurus.

XX

Key Location/Qualifiers

FT CDS 1071..2269

FT /tag= a

FT /product= "ScR/avidin fusion protein"

XX

PN WO9942577-A2.

XX

PD 26-AUG-1999.

XX

PF 23-FEB-1999; 99WO-GB0000546.

XX

PR 23-FEB-1998; 98GB-00003757.
 PR 24-JUN-1998; 98GB-00013653.
 XX (EURO-) EUROGENE LTD.
 XX Yla-Herttua S, Kulomaa M, Lehtolainen P, Marjomaki V, Airenne K;
 FI WPI; 1999-561345/47.
 DR P-PSDB; AAY03098.
 DR
 XX New fusion proteins having an extracellular domain with biotin-binding
 PT activity, used to target biotinylated molecules to specific sites in
 PT tissues.
 XX
 PS Disclosure; Page 15-21; 23pp; English.
 XX
 CC This invention describes a novel protein (A) which comprises a membrane-
 CC spanning domain and an extracellular domain (ECD), where the ECD
 CC comprises biotin-binding activity. Using the proteins or encoding nucleic
 CC acid molecules it is possible to target biotinylated molecules to
 CC specific sites in tissues. Molecules targeted in this way may be taken up
 CC by the tissues or cells by endocytosis, allowing the molecules to exert
 CC their effects within or on the cell. This sequence encodes a bovine
 CC scavenger receptor class A/avidin fusion protein which is used in the
 CC description of the invention
 XX
 SQ Sequence 5177 BP; 1356 A; 1222 C; 1390 G; 1209 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3.86e-177 Length: 5177
 Score: 2059.00 Matches: 400
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 2 Gaps: 0

US-10-618-570-2 (1-400) x AAZ09997 (1-5177)

QY 1 MetAlaGlnTrpAspAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
 DB 1071 ATGGCACAGTGGGATGACTTTCTCGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCT 1130
 QY 21 ValLysPheAspAlaAAspSerValThrAlaLeuLeuProProHisProLysAsnGlyPro 40
 DB 1131 GTGAAGTTCGATGCTCGCTCAGTGACAGCTTTGCTCTCCCTCCCATCTCTAAANAATGGCCCA 1190
 QY 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuLysLeuVal 60
 DB 1191 ACTCTTCAAGAGAGGATGAAGTCTTATAAACTGCACCTGATCACCTTTATCTCATTTGTG 1250
 QY 61 PheValValLeuValProIleGlyIleValAlaGlnLeuLeuLysTrpGluThr 80
 DB 1251 TTTGTAGTTCCTCGTCCCATCATTTGGCATATGTCAGCTCAGCTCCTGAAATGGGAACG 1310
 QY 81 LysAsnCysThrValGlySerValAsnAlaAspIleSerProSerProGluGlyLysGly 100
 DB 1311 AAGAATTCACGGTGGCTCGATTGATGCAAGATATCTCCAGTCCGGAAGGCAAGA 1370
 QY 101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120
 DB 1371 AATGCGAGTGAAGATGAATGAGATTTTCGAGAGCTGTGTGGAACGCATGAGCAACATG 1430
 QY 121 GluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPhe 140
 DB 1431 GAAAGCAGAAATCCAGTATCTTTTCAGATAATGAAGCCATCTCCTAGATGCTTAAGAATTC 1490
 QY 141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
 DB 1491 CAANAATTCACATAACAACTGATCAAGAATTTAATGATGTTCTTTCCAGCTAAATTC 1550
 QY 161 LeuLeuSerSerIleGlnGluHisGluAsnIleLeuGlyAspIleSerLysSerLeuVal 180
 DB 1551 TTACTTTCTCCATCCAGGAACATGAGAAATATCATAGGGGATATCTCCAAAGTCAATTAGTA 1610

QY 181 GlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal 200
 DB 1611 GGTCTGAACACCACAGTACTTTGATTTTCAGTTCTAGTATTGAAACACTGAATGGCAGAGTC 1670
 QY 201 GlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyrAsn 220
 DB 1671 CAAGAGAAATGCATTTTAAACAACAAGAGAGATGCGTAAATTTAGAGGCGGTATATACAAT 1730
 QY 221 AlaSerAlaGluLeuLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluLys 240
 DB 1731 GCATCAGCAGAAATTAAGTCTCTAGATGAAACCAAGTATATTTTGGAAACAGGAATAAAA 1790
 QY 241 GlyGluMetLysLeuLeuAsnIleThrAsnAspLeuArgLeuLysAspTrpGluHis 260
 DB 1791 GGGGAAATGAACACTGTTGAATAATATCACTAATGATCTGAGGCTGAAGAGTTGGAAACAT 1850
 QY 261 SerGlnThrLeuLysAsnIleThrLeuLeuGlnGlyAlaArgLysCysSerLeuThrGly 280
 DB 1851 TCTCAGACATTTGAAATAATATCACTTTACTCCAAGTCCAGAAAAGTGTCTGCTGACTGGG 1910
 QY 281 LysTrpThrAsnAspLeuGlySerAsnMetThrIleGlyAlaValAsnSerArgGlyGlu 300
 DB 1911 AAATGGACCAACGATCTGGGCTCCACATGACCATCGGGCTGTGAAACAGCAGAGGTGAA 1970
 QY 301 PheThrGlyThrTyrIleThrAlaValThrAlaThrSerAsnGluIleLysGluSerPro 320
 DB 1971 TTACAGGCACCTACATCAGCGGTAAACGCCACATCAANTAGATCAAAAGATCACCA 2030
 QY 321 LeuHisGlyThrGlnAsnThrIleAsnLysArgThrGlnProThrPheGlyPheThrVal 340
 DB 2031 CTGCATGGGACACAAACACCATCAACAAGAGGACCCAGCCACCTTTGGCTTCCACCGTC 2090
 QY 341 AsnTrpLysPheSerGluSerThrThrValPheThrGlyGlnCysPheIleAspArgAsn 360
 DB 2091 AATTGGAAGTTTTCAGAGTCCACCATCTCTTCACGGGCCAGTGTCTTCATAGACAGGAAT 2150
 QY 361 GlyLysGluValLeuLysThrMetTrpLeuLeuArgSerValAsnAspIleGlyAsp 380
 DB 2151 GGGAAAGAGGTCTCGAAGACCATGTGGCTCTGCGGTCAAGTGTAAATGACATTTGGTGTAT 2210
 QY 381 AspTrpLysAlaThrArgValGlyIleAsnIlePheThrArgLeuArgThrGlnLysGlu 400
 DB 2211 GACTGGAAAGCTACCAGGTCCGCATCAACATCTTCACTCGCTCGCCGACACACAGAGGAG 2270

RESULT 2
 AAQ04926
 ID AAQ04926 standard; DNA; 1588 BP.
 XX
 AC AAQ04926;
 XX
 DT 25-MAR-2003 (revised)
 DT 31-OCT-2002 (revised)
 DT 23-OCT-1990 (first entry)
 XX
 DE Sequence encoding portion of scavenger receptor protein with affinity for
 DE acylated low density lipoprotein (aLDL).
 XX
 XX Acylated low density lipoprotein; aLDL; atherosclerotic plaque; ds.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 4..1362
 FT /*tag= a
 XX
 XX WO9005748-A.
 XX
 PD 31-MAY-1990.
 XX
 PF 15-NOV-1988; 88US-00272002.
 XX
 PR 15-NOV-1988; 88US-00272002..

484	Db	TTACTTCTCCATCCAGGAACATGAGATATATCATAGGGGATATCTCCAAAGTCATTAGTA	543
181	Qy	GlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal	200
544	Db	GGTCTGAACACCCACAGTACTTGATTTGCAGTTTCAGTATTGAAACACTGATGCGAGATC	603
201	Qy	GlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyrAsn	220
604	Db	CAAGAGAATGCATTTAAACAACAAGAGGAGATCGCTAAATTTAGAGGAGCGTATATACAAT	663
221	Qy	AlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIleLys	240
664	Db	GCATCAGCAAAATTAAGTCTCTAGATGAAAAACAAGTATATTTTGGAAACAGGAAATAAAA	723
241	Qy	GlyGluMetLysLeuLeuAsnAenIleThrAsnAspLeuArgLeuLysAspTrpGluHis	260
724	Db	GGGGAATGGAACTGTTGAATAATATCACTAATGATCTGAGGCTGAAGGATTTGGGAACAT	783
261	Qy	SerGlnThrLeuLysAsnIleThrLeuGlnGlnGlyAla-----	273
784	Db	TCTCAGACATGAAAAAATATCACTTTACTCCAAGTCTCTCGAGGCTCCAGGTGAAAAA	843
274	Qy	-----ArgLysCysSerLeuThrGlyLysTyrThrAsnAspLeuGly---SerAsnMet	290
844	Db	GGAGATAGAGGCCCTCCCTGGACA---AAATGGTATATCCAGGCTTTCCAGGCTCTAATAGG	899
291	Qy	ThrIleGlyAlaValAsnSerArgGlyGluPheThrGlyThrTyrIleThrAlaValThr	310
900	Db	TACTCCAGGCTCTAAAGGTGATCGGGGATCTCTGGTTTACCTGGAGTTCGAGGATCCCC	959
311	Qy	AlaThrSerAsnGluLysGluSerProLeuHisGlyThrGlnAsnThrIleAsnLys	330
960	Db	AGGACCAATGGGGAACACCGGGAAGCCAGCTTAAATGGACAAAAAAGGCCAGAAAGGAGA	1019
330	Qy	sArgThrGlnProThrPhe-----GlyPheThrVal-AsnTrpLysPheSerGluS	347
1020	Db	AAAAACCATCCCAACCATCCAAACACAACTAAATACAGTCCGACTGGTGGGTGGAGCGG	1079
347	Qy	erThrThrValPheThrGlyGlnCysPheIleAspArgAsnGlyLysGluValLeuLysT	367
1080	Db	CCCTCAGAAGGCAGAGTGGAGATTTTTCACGAGGCCAGTGGGGTACGGTGTGTGACGA	1139
367	Qy	hrMetTrpLeuLeuArgSerSerVal-AsnAspIleGlyAsp---AspTrpLysAlaThr	385
1140	Db	CCGC-TGGGAACCTGCGTGGAGGACTCGTCTGTCGAGGAGCTTGGGATACAAAGGTGTC	1198
386	Qy	ArgValGlyIleAsn-----IlePheThrArgLeuArgThrGln	398
1199	Db	AAAGTGTGCATAAGCAGCTTATTTTGGAAAAAGGTACGGGTCCAA	1243
RESULT 4			
AAQ04925			
ID	AAQ04925 standard; DNA; 1807 BP.		
XX	AAQ04925;		
XX			
DT	25-MAR-2003 (revised)		
DT	31-OCT-2002 (revised)		
DT	23-OCT-1990 (first entry)		
XX			
DE	Sequence encoding scavenger receptor protein with affinity for acylated		
DE	low density lipoprotein (aLDL).		
XX			
KW	Acylated low density lipoprotein; aLDL; atherosclerotic plaque; ds.		
XX			
OS	Synthetic.		
XX			
FH	Key Location/Qualifiers		
FT	36..1082		
FT	/*tag= a		
XX			
PN	WO9005748-A.		
XX			

Alignment Scores:			
Pred. No.:	4,78e-116	Length:	1588
Score:	1381.50	Matches:	307
Percent Similarity:	79.4%	Conservative:	24
Best Local Similarity:	73.6%	Mismatches:	66
Query Match:	67.1%	Indels:	22
DB:	2	Gaps:	6
US-10-618-570-2 (1-400) x AAQ28540 (1-1588)			
Qy	1	MetAlaGlnTTrpAspAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer	20
Db	4	ATGGCACAGTGGCATGACTTTCTTGATCAGCAAGAGGACACTGCACAGCTGTACAGAGTCT	63
Qy	21	ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro	40
Db	64	GTAAGTTGATGCTCGCTCAGTCACAGCTTTGCTTCTCCCATCTCTAAAATGGCCCA	123
Qy	41	ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal	60
Db	124	ACTCTTCAAGAGAGGATGAAGTCTTATAAAACTGCACCTGATCACCTTTATCTCATTTGTG	183
Qy	61	PheValValLeuValProIleGlyIleValalaGlnLeuLeuLysTyrGluThr	80
Db	184	TTTGATGTTCTCGTGCCCATCATTTGGCATATGTCAGCTCAGCTCTGAAATGGGAACG	243
Qy	81	LysAsnCysThrValGlySerValAsnAlaAspIleSerProSerProGluGlyLYBSGly	100
Db	244	AAGAATTGCACGGTTGGCTCAGTTAAATGCAGATATATCTCCAAGTCGGGAAGGCAAAAG	303
Qy	101	AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet	120
Db	304	AATGGCAGTGAAGATGAATGAGATTTTCGAGAAGCTGTGTATGGAACGCATGAGCAACATG	363
Qy	121	GluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPhe	140
Db	364	GAAAGCAGAAATCCAGATATCTTTTCAGATAATCAAGCCCAATCTCTTAGATGCTAAGAATTC	423
Qy	141	GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer	160
Db	424	CAAAATTTTCAGCATAAACACTGATCAAGATTTTAATGATGTCTTTTCCAGCTCAAAATTC	483
Qy	161	LeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeuVal	180

PD 31-MAY-1990.
 XX 15-NOV-1988; 88US-00272002.
 XX 15-NOV-1988; 88US-00272002.
 PR 09-AUG-1989; 89US-00391486.
 XX (MASI) MASSACHUSETTS INST TECHNOLOGY.
 XX Kodama T, Krieger M;
 PI WPI; 1990-193408/25.
 XX P-PSDB; AAR05508.
 DR
 DR
 XX
 PT New receptor protein - has affinity for acetylated low density
 PT lipoprotein and corresponding antibodies and DNA sequences.
 XX
 XX Disclosure; Page ?; 79pp; English.
 PS
 CC Receptor protein, and fragments and analogues thereof may be immobilised
 CC on a support and used in assay and purification of the aLDL target.
 CC Labelled Abs, raised to the protein may be injected into the vascular
 CC system to detect the presence of atherosclerotic plaques. (Updated on 31-
 CC OCT-2002 to add missing OS field.) (Updated on 25-MAR-2003 to correct PR
 CC field.) (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-
 CC 2003 to correct PI field.)
 XX
 XX Sequence 1807 BP; 556 A; 384 C; 406 G; 461 T; 0 U; 0 Other;
 SQ

Alignment Scores:
 Pred. No.: 5.63e-116 Length: 1807
 Score: 1381.50 Matches: 288
 Percent Similarity: 85.2% Conservative: 16
 Best Local Similarity: 80.7% Mismatches: 34
 Query Match: 67.1% Indels: 20
 Gaps: 4

US-10-618-570-2 (1-400) x ARA04925 (1-1807)
 QY 1 MetAlaGlnTrpAspPheProAspGlnGluAspThrAspSerCysThrGluSer 20
 DB 36 ATGCACAGTGGGATGACTTTCCTGATCAGCAAGAGCACTGACAGCTGTACAGATCT 95
 QY 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
 DB 96 GTGAAGTTCGATGCTCGCTCAGTCAGCAGCTTGTCTCTCCCATCTTAAATGGCCCA 155
 QY 41 ThrLeuGlnGluArgMetLysSerTyrlsThrAlaLeuLeuThrLeuTyrlsLeuVal 60
 DB 156 ACTCTTCAAGAGAGGATGAAGTCTTATAAACTGCACCTGATCACCCCTTATCTCATTTGTG 215
 QY 61 PheValValLeuValProIleileGlyLeValAlaAlaGlnLeuLeuLysTrpGluThr 80
 DB 216 TTGTGTGTTCTGGTCCCATCTTGGCAGTATGTCAGCTCAGCTCCTGAAATGGGAACG 275
 QY 81 LysAsnCysThrValGlySerValAsnAlaAspIleSerProSerProGluGlyLysGly 100
 DB 276 AAGAATTGCACGGTTGGCTCAGTTAATGCAGATATATCTCCAGTCCGGAAGGCAAGGA 335
 QY 101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120
 DB 336 AATGGCAGTGAAGATGAATGAGATTTTCGAGAAGCTGTGTATGGAACGCGATGAGCAACATG 395
 QY 121 GluSerArgIleGlnTyrlsLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPhe 140
 DB 396 GAAAGCAGATCCAGTATCTTTCAGATATGAAGCCATCTCTAGATGCTTAAGATTTTC 455
 QY 141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
 DB 456 CAAATTTTCAGCATPAACAACCTGATCAAGATTTAATGATGTTCTTTCCAGCTAAATTC 515
 QY 161 LeuLeuSerSerIleGlnGluHisGluAsnIleleGlyAspIleSerLysSerLeuVal 180

Db 516 TTACTTTCTCCATCCAGGAACATGAGAAATATCATAGGGATATCTCCAAGTCATTAGTA 575
 QY 181 GlyLeuAsnThrThrValLeuLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal 200
 Db 576 GGTCTGAACACACACAGTACTTGTATTTGCAGTTTCAGTATTGAAACACCTGAATGGCAGAGTC 635
 QY 201 GlnGluAsnAlaPheLysGlnGlnGluGluMetArgLysLeuGluGluArgIleTyrlsAsn 220
 Db 636 CAAGAGAAATGCATTTAAACCAACAGAGGAGATGCGTAAATTTAGAGGAGCGGTATATACAAT 695
 QY 221 AlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrlsLeuGlnGlnGluIleLys 240
 Db 696 GCATCAGCAGAAATTAAGTCTCTAGATGAATAACAGTATATTTGGACAGGAATATAAA 755
 QY 241 GlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGluHis 260
 Db 756 GGGGAATGAACACTGTTGAATAATATCACTAATGATCTGAGGCTGAAGGATTGGGAACAT 815
 QY 261 SerGlnThrLeuLysAsnIleThrLeuLeuGlnGlyAla----- 273
 Db 816 TCTCAGACATTTGAAAAATATCACTTTACTCAAGGCTCTCTCGACCTCCAGGTGAAAAA 875
 QY 274 -----ArgLysCysSerLeuThrGlyLysTrpThrAsnAspLeuGly---SerAsnMet 290
 Db 876 GGAGATAGAGGCC-TCTCGACA---AATGGTATACCAAGGCTTTCAGGTCTAATAGG 931
 QY 291 ThrIleGlyAlaValAsnSerArgGlyGluPheThrGlyThrTyrlsThrAlaValThr 310
 Db 932 TACTCCAGGCTTAAAGGTGATCGGGGATCTCTGGTTTACCTGGAGTTCGAGGATTCCTCC 991
 QY 311 AlaThrSerAsnGluIleLysGluSerProLeuHisGlyThrGlnAsnThrIleAsnLys 330
 Db 992 AGGACCAATGGGAAGACCGGAAGCCAGGACTTATTGGCAAAAGCCAGAGGAGAGA 1051
 QY 331 ArgThrGlnProThrPheGlyPheThrValAsnTrpLysPheSerGluSer 347
 Db 1052 ARAAGGAG-----TGAAGCATGCAAGACC 1078

RESULT 5
 ADB89012
 ID ADB89012 standard; DNA; 1077 BP.
 XX
 AC ADB89012;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Human DNA encoding the macrophage scavenging receptor.
 XX
 KW Human; db; macrophage scavenger receptor; heptad repeat;
 KW reduced immunogenicity; trimerisation; CD40L; immune response;
 KW T-cell activation; isotype switching; gene.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 1..1077
 FT /*tag= a
 FT /product= "Macrophage scavenger receptor"
 XX
 FN US2003119149-A1.
 XX
 XX 26-JUN-2003.
 XX
 XX 20-DEC-2002; 2002US-00326186.
 PF
 PF 21-DEC-2001; 2001US-0343315P.
 PR
 XX (REDD/) REDDY P.
 XX
 XX Reddy P;
 XX
 DR WPI; 2003-863437/80.
 DR P-PSDB; ADB89003.

XX New fusion polypeptide that is capable of forming a trimer, useful in
 PT preparing a composition for treating immune related or inflammatory
 PT disorders.
 XX
 PS Example; Fig 4; 24pp; English.
 XX
 CC The present invention relates to a polypeptide comprising a fusion
 CC protein capable of forming a trimer and comprising a heptad repeat
 CC containing region of the human macrophage scavenging factor protein
 CC appearing as ADB89007 fused to a second protein (especially a member of
 CC the tumour necrosis superfamily, TNFSF e.g. CD40L). CD40L is involved in
 CC isotype switching and is important in T-cell activation in the immune
 CC response. The fusion protein has reduced immunogenicity due to its
 CC ability to form trimers. Also included are the nucleic acids encoding the
 CC fusion proteins, host cells comprising the nucleic acids (used to produce
 CC the fusion protein) and a recombinant vector comprising the nucleic acid.
 CC The fusion proteins are used to deliver a therapeutic protein with
 CC reduced immunogenicity to a patient in need of the protein. The present
 CC sequence encodes the human macrophage scavenging factor protein. Note:
 CC The authors have mixed up the descriptions of the sequences as they
 CC appear in the figures, therefore there is a possibility that the
 CC description assigned by the indexer may be wrong.
 XX

SQ Sequence 1077 BP; 364 A; 221 C; 242 G; 250 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2,29e-85 Length: 1077
 Score: 1042.50 Matches: 197
 Percent Similarity: 87.2% Conservative: 41
 Best Local Similarity: 72.2% Mismatches: 34
 Query Match: 50.6% Indels: 1
 DB: 10 Gaps: 1

US-10-618-570-2 (1-400) x ADB89012 (1-1077)

Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
 Db 1 ATGGAGCAGTGGGATCACTTTTCACATCAACAGAGGACACTGATAGCTGCTCGAATCT 60
 Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
 Db 61 GTGAAATTTGATGCTCGCTCAATGACAGCTTTGCTTCTCCGATCTTAAACACAGCCCT 120
 Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
 Db 121 TCCCTTCAAGAGAAACTGAGCTCTTCAAGCTGCACTGATTCCTCTTACCTCTCGTG 180
 Qy 61 PheValValLeuValProIleGlyLeValAlaAlaGlnLeuLeuTyrGluThr 80
 Db 181 TTTCAGTTCATCCCTCTCATTTGGAATAGTGGAGCTCAACTCTCGAAGTGGGAACG 240
 Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
 Db 241 AAGAATTGCTCAGTTAGTTCAACTAATGCAATATGATATATACTCAAGTCTCACGGGAAA 300
 Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
 Db 301 GGAATGACAGCGAAGAGAAATGAGATTTTCAAGAAAGTCTTTATGGAACACATCAGCAAC 360
 Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
 Db 361 ATGGAGAAGAAATCCAGCATATTTTAGCATGGAAGCCACCTCATGGACACAGAGCAT 420
 Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
 Db 421 TTCCAAAATTTTCAGCATGACAACTGATCAAGATTTAATGACATTTCTGCACCTAAGT 480
 Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleLeuGlyAspIleSerLysSerLeu 179
 Db 481 ACCTTGTTTCTCAGTCAGGACATGGGAATGCAATAGATGAATCTCCAACTCCTTA 540
 Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199

Db 541 ATAAGTTTGATACCACTGCTTGATTGCGCTCAACATAGAAATCTGAATGGCAA 600
 Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
 Db 601 ATCCAGAGAATACCTTCAAAACAACAAGAGAAATCAGTAAATTAGAGGACGCTGTTTAC 660
 Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
 Db 661 AATGTATATCAGCAGAAATATATGGCTATGAAAGAGAAACAAGTGCATTTGGAAACAGGAATA 720
 Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTyrGlu 259
 Db 721 AAGGAGAGAGTGAAGTACTGATTAACATCACTAATGATCTCAGACTGAAAGATTGGAA 780
 Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
 Db 781 CATTCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 819

RESULT 6

AAQ21550
 ID AAQ21550 standard; DNA; 1347 BP.

XX AAQ21550;

XX 25-MAR-2003 (revised)

DT 19-MAY-1992 (first entry)

XX Gene encoding human scavenger receptor type II.

DE Lipoproteins; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FF CDS 47..1124

FT /*tag= a

FT /product= "scavenger_receptor_type_II"

XX JF03290184-A.

XX 19-DEC-1991.

XX 06-APR-1990; 90JP-00090274.

XX 06-APR-1990; 90JP-00090274.

XX (CHUS) CHUGAI PHARM CO LTD.

XX WPI; 1992-051436/07.

DR P-PSDB; AAR21513.

XX New scavenger receptor-producing animal cells - which have been
 PT transformed with vector contg. gene for coding human scavenger receptor I
 PT or II type under control of promoter, etc.

XX Disclosure; Fig 2; 9pp; Japanese.

XX The gene can be expressed from a cytomegalovirus promoter in a host cell
 CC e.g. CHO cell. The resulting recombinant scavenger receptor can be used
 CC to detect modified lipoproteins or modified substances in the blood. See
 CC also AAQ21549. (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 1347 BP; 420 A; 297 C; 298 G; 332 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3,05e-85 Length: 1347
 Score: 1042.50 Matches: 197
 Percent Similarity: 87.2% Conservative: 41
 Best Local Similarity: 72.2% Mismatches: 34
 Query Match: 50.6% Indels: 1
 DB: 2 Gaps: 1


```

Db 467 TTTCAAAATTTTCAGCATGACAACTGATCAAGATTTAATGACATTTCTCTCGAGCTAAGT 526
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAenIleIleGlyAepIleSerIleSerLeu 179
Db 527 ACCTTTGTTTCTCAGTCCAGGACATGGGAATGCAATAGATGAATCTCCAAGTCTCTTA 586
Qy 180 ValGlyLeuAenThrThrValLeuAepLeuGlnPheSerIleGluThrLeuAenGlyArg 199
Db 587 ATAAGTTTGAATACCATGCTTGAATGTTGAGCTCAACATAGAAATCTGAATGGCAAA 646
Qy 200 ValGlnGluAenAlaPheLysGlnGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 647 ATCCAAGAGAATACCTTCAACAACAAGAGAAATCAGTAAATTAGAGAGCGTGTTTAC 706
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAepGluLysGlnValTyrLeuGluGlnGluIle 239
Db 707 AATGTATCAGCAGAAATTTATGCTATGAAAGAAAGAACCAAGTGCATTTGGAAACAGAAATA 766
Qy 240 LysGlyGluMetLysLeuAenAenIleThrAenAepLeuArgLeuAenThrGlu 259
Db 767 AAGGAGAAGTGAAGTACTGAATACATCACTAATGATCTCAGACTGAAGATTTGGAA 826
Qy 260 HisSerGlnThrLeuLysAenIleThrLeuLeuGlnGly 272
Db 827 CATCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 865

RESULT 8
AAF29165
ID AAF29165 standard; DNA; 1347 BP.
AC AAF29165;
XX
XX
DT 05-APR-2001 (first entry)
DE Human DNA encoding type II scavenger receptor.
XX
XX Human; type II scavenger receptor; antibody; macrophage; ds.
XX
XX Homo sapiens.
XX
PN JP2000312595-A.
PD 14-NOV-2000.
XX
PF 16-AUG-1991; 2000JP-00114233.
XX
XX 27-AUG-1990; 90JP-00222398.
PR 16-AUG-1991; 91JP-00229728.
XX
XX (CHUS ) CHUGAI PHARM CO LTD.
XX
XX WPI; 2001-161981/17.
DR P-PSDB; AAB49716.
XX
XX New anti-human scavenger receptor antibody useful for the identification
PT and the determination of macrophages.
XX
XX Example 1; Page 7-9; 13pp; Japanese.
XX
XX This invention relates to a poly or monoclonal antibody which targets a
CC human type I or type II scavenger receptor epitope. The anti-scavenger
CC receptor antibody can be used for the identification and determination of
CC macrophages. The present sequence represents DNA encoding the human type
CC II scavenger receptor
XX
XX Sequence 1347 BP; 420 A; 297 C; 298 G; 332 T; 0 U; 0 Other;

```

Alignment Scores:

```

Pred. No.: 3.05e-85 Length: 1347
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34

```

```

Query Match: 50.6% Indels: 1
DB: 4 Gaps: 1

US-10-618-570-2 (1-400) x AAF29165 (1-1347)

Qy 1 MetAlaGlnTrpAepAepPheProAepGlnGlnGlnAepThrAepSerCyThrGluSer 20
Db 47 ATGGAGCAGTGGGATCACTTTCACAATCAACAGAGGACACTGATAGCTGCTCCGAATCT 106
Qy 21 ValLysPheAepAlaArgSerValThrAlaLeuLeuProProHisProLysAenGlyPro 40
Db 107 GTGAAATTTGATGCTCGCTCAATGACAGCTTTGCTTCTCCGAATCTCTAAACACAGCCCT 166
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60
Db 167 TCCCTTCACAGAGAACTGAAAGTCTTCAAGCTGCATGATTGCCCTTTACTCTCTCGTG 226
Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTyrGluThr 80
Db 227 TTTGCAGTTCATCCCTCTCATTTGGAATAGTGCAGCTCAACTTCTTGAAGTGGAAACG 286
Qy 81 LysAenCysThrValGlySerValAenAla---AspIleSerProSerProGluGlyLys 99
Db 287 AAGAAATTCGTAGTTCAGTTCCTCAACTTAATGCAATGATAAATCAAGTCTCACGGGAAA 346
Qy 100 GlyAenGlySerGluAepGluMetArgPheArgGluAlaValMetGluArgMetSerAen 119
Db 347 GGAATGACAGCGAAGAGGAATGAGATTTCAAGAGCTCTTTATGGAAACATCATGAGCAAC 406
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAenGluAlaAenLeuLeuAenAlaLysAen 139
Db 407 ATGGAGAAGAGATCCAGCATATTTTAGACATGGAAGCCCAACCTCATGGACACAGAGCAT 466
Qy 140 PheGlnAenPheSerIleThrThrAepGlnArgPheAenAepValLeuPheGlnLeuAen 159
Db 467 TTCCAAATTTTCAGCATGACAACTGATCAAGATTTAATGACATTTCTTCTGCAGTAAGT 526
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAenIleIleGlyAepIleSerLysSerLeu 179
Db 527 ACCTTTGTTTCTCAGTCCAGGACATGGGAATGCAATAGATGAATCTCCAGTCTCTTA 586
Qy 180 ValGlyLeuAenThrThrValLeuAepLeuGlnPheSerIleGluThrLeuAenGlyArg 199
Db 587 ATAAGTTTGAATACCACTTGTGATTTCAGCTCAACATAGAAATCTGATGGCAAA 646
Qy 200 ValGlnGluAenAlaPheLysGlnGlnGlnGluMetArgLysLeuGluArgIleTyr 219
Db 647 ATCCACAGAGAATACCTTCAACAACAAGAGGAAATCAGTAAATTAGAGAGCGTGTTTAC 706
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAepGluLysGlnValTyrLeuGluGlnGluIle 239
Db 707 AATGTATCAGCAGAAATTTATGCTATGAAAGAAAGAACCAAGTGCATTTGGAAACAGAAATA 766
Qy 240 LysGlyGluMetLysLeuLeuAenAenIleThrAenAepLeuArgLeuLysAepTyrGlu 259
Db 767 AAGGAGAAGTGAAGTACTGAATACATCACTAATGATCTCAGACTGAAGATTTGGAA 826
Qy 260 HisSerGlnThrLeuLysAenIleThrLeuLeuGlnGly 272
Db 827 CATCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 865

RESULT 9
AAT43260
ID AAT43260 standard; cDNA to mRNA; 1367 BP.
XX
XX AAT43260;
XX
XX AC
XX
XX 25-FEB-1997 (first entry)
XX
XX Type II macrophage scavenger receptor gene.
XX
XX Type II macrophage scavenger receptor; SRG; kidney; 293 cell;
XX cell substrate; cell attachment; atherosclerosis; ds.

```


CC discovery and development, toxicological and carcinogenicity studies,
 CC forensics and pharmacogenomics. The microarray is also useful for
 CC monitoring progression of diseases and for developing sophisticated
 CC profiles for the effects of currently available therapeutic drugs. The
 CC combination is also useful for purifying a subpopulation of mRNAs, cDNAs
 CC and genomic fragments and in research and diagnostic applications. The
 CC array can detect changes in expression in a large number of genes coding
 CC for different signaling pathway populations which can be used to diagnose
 CC various diseases including cancer e.g. adenocarcinoma and leukemia,
 CC immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease
 CC and Parkinson's disease. The present sequence represents a polynucleotide
 CC probe of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=0650093881

XX Sequence 1367 BP; 427 A; 298 C; 303 G; 339 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3,11e-85 Length: 1367
 Score: 1042.50 Matches: 197
 Percent Similarity: 87.2% Conservatives: 41
 Best Local Similarity: 72.2% Mismatches: 34
 Query Match: 50.6% Indels: 1
 DB: 10 Gaps: 1

US-10-618-570-2 (1-400) x ACA56662 (1-1367)

Qy 1 MetAlaGlnTrrAppAppPheProAspGlnGlnGlnAspThrAspSerCysThrGluSer 20
 Db 67 ATGGAGCAGTGGGATCACTTTTCACAATCAACAGGAGGACACTGATAGCTGCCGAATCT 126
 Qy 21 ValLysPheAepAlaAqgSerValThrAlaLeuLeuProHisProLysAenGlyPro 40
 Db 127 GTGAAATTGATGCTCGCTCAATGACAGCTTGTCTCTCGAATCTTAAACACAGCCCT 186
 Qy 41 ThrLeuGlnGluArgMetLysSerTyxThrAlaLeuLeuThrLeuTyrLeuIleVal 60
 Db 187 TCCCTTCAAGAGAAGAACTGAGTCTCTCAAGCTGCACGTGTCCTTACCTCTCTGTG 246
 Qy 61 PheValValLeuValProIleGlyLeValAlaAlaGlnLeuLeuLysTrrPgluThr 80
 Db 247 TTTGAGTTCTCTCATCTCTCTCATTTGGAATAGTGGCAGCTCAACTCTCTGAAAGTGGAAACG 306
 Qy 81 LysAenCysThrValGlySerValAsnAla---AspLysSerProSerProGluGlyLys 99
 Db 307 AAGAAATTGCTCAGTTAGTTCAACTAATGCAATGATATACTCAAGTCTCACGGGAAA 366
 Qy 100 GlyAenGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
 Db 367 GGAATGCACAGCGAAGAGGAATGAGATTTCAAGAGTCTTTATGGACACATGAGCAAC 426
 Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAenLeuLeuAspAlaLysAen 139
 Db 427 ATGGAGAAGAAATCCAGCATATTTAGACATGGAAGCCAACTCATGGACACAGAGCAT 486
 Qy 140 PheGlnAenPheSerIleThrAspGlnArgPheAenAspValLeuPheGlnLeuAen 159
 Db 487 TTTCAAAATTTTCAGCATGACAACTGATCAAGATTTTAAATGACATCTTCTGACACTAAGT 546
 Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAenIleIleGlyAspIleSerLysSerLeu 179
 Db 547 ACCTTGTCTTCTCAGTCCAGGACATGGGATGCAATAGATGAATCTCCAGTCTCTTA 606
 Qy 180 ValGlyLeuAenThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAenGlyArg 199
 Db 607 ATAAGTTGAATACCACTGCTTGTGATTTGAGCTCAACATAGAAAATCTGAATGGCAAA 666
 Qy 200 ValGlnGluAenAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
 Db 667 ATCCAGAGAATACCTTCAAAACAAAGAGAAATCAGTAAATTTAGAGGACGCTTTAC 726
 Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239

Db 727 AATGATATCAGCAGAAATTTATGGCTATGAAAGAAGAACAGTGCATTTTGGACAGAAATA 786
 Qy 240 LysGlyGluMetLysLeuLeuAenAsnIleThrAsnAspLeuArgLeuLysAspTrrPglu 259
 Db 787 AAAGGAGAAGTGAAGTACTGAATAACATCATCACTAATGATCTCAGACTGAAAGATTGGNA 846
 Qy 260 HisSerGlnThrLeuLysAenIleThrLeuLeuGlnGly 272
 Db 847 CATTCAGACCTTGAGAAATATCATCTTTAATTCAGGT 885

RESULT 11

ADI56458
 ID ADI56458 standard; DNA; 1367 BP.

XX ADI56458;

XX 22-APR-2004 (first entry)

XX Human polynucleotide probe #1260.

XX Human; probe; ss; receptor-like polypeptide; transducing polypeptide;
 KW effector-like polypeptide; cancer; immunopathology; neuropathology;
 KW drug development; toxicology; carcinogenicity;
 KW signalling pathway polypeptide; adrenal gland; bladder; bone;
 KW bone marrow; brain; breast; cervix; tumour; immunopathology; AIDS;
 KW diabetes; pancreatitis; osteoporosis; ulcerative colitis; neuropathology;
 KW dementia; amnesia; epilepsy; Alzheimer's disease; depression.

XX Homo sapiens.

XX US2004010136-A1.

XX 15-JAN-2004.

XX 26-NOV-2002; 2002US-00305720.

XX 30-JAN-1998; 98US-00016434.

XX (INCY-) INCYTE GENOMICS INC.

XX Au-Young J, Seilhamer JJ;

XX WPI; 2004-090520/09.

XX New composition comprising polynucleotide probes, useful as array
 PT elements in a microarray for monitoring the expression of target
 PT polynucleotides or purifying a subpopulation of mRNAs, cDNA, or genomic
 PT fragments.

XX Claim 6; SEQ ID NO 1260; 73pp; English.

XX The invention relates to a composition of polynucleotide probes
 CC comprising first polynucleotide probes comprising at least a portion of a
 CC gene encoding a receptor-like polypeptide, second polynucleotide probes
 CC comprising at least a portion of a gene encoding a transducing
 CC polypeptide and third polynucleotide probes comprising at least a portion
 CC of a gene encoding an effector-like polypeptide. The probes of the
 CC composition are useful as array elements in a microarray for monitoring
 CC the expression of target polynucleotides. The microarray is useful in the
 CC diagnosis and treatment of cancer, an immunopathology or a
 CC neuropathology. It can also be used for drug discovery and development,
 CC toxicological and carcinogenicity studies, forensics or pharmacogenomics.
 CC Microarrays can also be used for monitoring the progression of diseases
 CC that may be associated with the altered expression of signalling pathway
 CC polypeptides. The composition can also be used to purify a subpopulation
 CC of mRNAs, cDNAs, or genomic fragments in a sample. The expression profile
 CC is also useful for the diagnosis and treatment of cancer, e.g. cancers of
 CC the adrenal gland, bladder, bone, bone marrow, brain, breast or cervix,
 CC an immunopathology, e.g. AIDS, diabetes, pancreatitis, osteoporosis or
 CC ulcerative colitis, or a neuropathology, e.g. dementia, amnesia,
 CC epilepsy, Alzheimer's disease or depression. This sequence represents a
 CC human polynucleotide probe of the invention. Note: The sequence data for

CC this patent did not form part of the printed specification but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html.

SQ Sequence 1367 BP; 427 A; 298 C; 303 G; 339 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3,11e-85 Length: 1367
 Score: 1042.50 Matches: 197
 Percent Similarity: 87.2% Conservative: 41
 Best Local Similarity: 72.2% Mismatches: 34
 Query Match: 50.6% Indels: 1
 DB: 12 Gaps: 1

US-10-618-570-2 (1-400) x ADI56458 (1-1367)

Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
 Db 67 ATGGAGCAGTGGGATCACTTTTCAATCAACAGAGGACACTGATAGTCTCGGATCT 126
 Qy 21 VallysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
 Db 127 GTGAATTTGATGCTCGCTCAATGACAGCTTTGCTTCTCGGAATCTTAAACACAGCCCT 186
 Qy 41 ThrLeuGlnGluArgMetLysSerTyrlsThrAlaLeuLeuThrLeuTyLeuVal 60
 Db 187 TCCCTTCAAGAGAACTCAAGTCTTCAAGCTGCACGTATGCCCTTTACCTCTCGTG 246
 Qy 61 PheValValLeuValProIleIleGlyLeValAlaAlaGlnLeuLeuTyTrpGluThr 80
 Db 247 TTTGAGTTCTCATCCCTCTCATTTGGATATGTCGAGCTCACTTCTGAAATGGGAACG 306
 Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
 Db 307 AAGAATTGCTCAGTTAGTTCAACTAATGCAATGATATACTCAAGTCTCACGGGAAA 366
 Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
 Db 367 GGAATGACACGGAAGAGGAATGAGATTTTCAAGAGTCTTTATGGAAACATGAGCAAC 426
 Qy 120 MetGluSerArgIleGlnTyLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
 Db 427 ATGGAGAAGAGATCCAGCATATTTAGCATGGAAGCCAACTCATGGAACAGAGCAT 486
 Qy 140 PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
 Db 487 TTCCAAATTTTCAGCATGACACTGATCAAGATTTAATGACATTTCTCGCAGTAAGT 546
 Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
 Db 547 ACCTTGTCTTCTCAGTCCAGGACATGGGAATGCAATAGATGAATCTCCAAGTCTCTTA 606
 Qy 180 ValGlyLeuAsnThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
 Db 607 ATAAGTTTGAATACCATCTGCTGATTTGAGCTCAACATAGAAATCTGAATGGCAAA 666
 Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTy 219
 Db 667 ATCCAGAGATACCTTCAACACACAGAGGAATCAGTAATTAGAGGACGCTGTATAC 726
 Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyLeuGluGluIle 239
 Db 727 AATGATATCAGCAGAAATATGCTATGAAAGAAACAAAGTGCATTTTGGAAACAGAAATA 786
 Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAsnTrpGlu 259
 Db 787 AAAGAGAAGTGAAGATGTAATCAACATCACTAATGATCTCAGACTGAAGATGGGAA 846
 Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
 Db 847 CATTTCTAGACCTTGAGAAATATCATCTTAATTCAGT 885

RESULT 12

AAQ21549
 ID AAQ21549 standard; DNA; 2028 BP.
 XX AC AAQ21549;
 XX 25-MAR-2003 (revised)
 DT 19-MAY-1992 (first entry)
 XX Gene encoding human scavenger receptor type I.
 DE Lipoproteins; ss.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 CDS 47..1402
 FT /*tag= a
 FT /product= "scavenger_receptor_type_I"
 FT
 FN JP03290184-A.
 XX 19-DEC-1991.
 XX 06-APR-1990; 90JP-00090274.
 XX 06-APR-1990; 90JP-00090274.
 XX (CHUS) CHUGAI PHARM CO LTD.
 XX WPI; 1992-051436/07.
 DR P-PSDB; AAR21512.
 XX
 PT New scavenger receptor-producing animal cells - which have been
 transformed with vector contg. gene for coding human scavenger receptor I
 or II type under control of promoter, etc.
 PS Disclosure; Fig 1; 9pp; Japanese.
 CC The gene can be expressed from a cytomegalovirus promoter in a host cell
 e.g. CHO cell. The resulting recombinant scavenger receptor can be used
 to detect modified lipoproteins or modified substances in the blood. See
 also AAQ21550. (Updated on 25-MAR-2003 to correct PA field.)
 SQ Sequence 2028 BP; 652 A; 369 C; 434 G; 573 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5,14e-85 Length: 2028
 Score: 1042.50 Matches: 197
 Percent Similarity: 87.2% Conservative: 41
 Best Local Similarity: 72.2% Mismatches: 34
 Query Match: 50.6% Indels: 1
 DB: 2 Gaps: 1
 US-10-618-570-2 (1-400) x AAQ21549 (1-2028)
 Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
 Db 47 ATGGAGCAGTGGGATCACTTTTCAATCAACAGAGGACACTGATAGTCTCGGATCT 106
 Qy 21 VallysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
 Db 107 GTGAATTTGATGCTCGCTCAATGACAGCTTTGCTTCTCGGAATCTTAAACACAGCCCT 166
 Qy 41 ThrLeuGlnGluArgMetLysSerTyrlsThrAlaLeuLeuThrLeuTyLeuVal 60
 Db 167 TCCCTTCAAGAGAACTCAAGTCTTCAAGCTGCACGTATGCCCTTTACCTCTCGTG 226
 Qy 61 PheValValLeuValProIleIleGlyLeValAlaAlaGlnLeuLeuTyTrpGluThr 80
 Db 227 TTTGAGTTCTCATCCCTCTCATTTGGATATGTCGAGCTCACTTCTGAAATGGGAACG 286
 Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99


```

|||||.....|||||.....|||||.....|||||.....|||||.....
767 AAAGGAGAGTGAAGTACTGATTAACATCACTAATGATCTCAGACTGAAAGATTGGAA 826
Db

260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
|||||.....|||||.....|||||.....|||||.....|||||.....
827 CATTTCTCAGAGCTTGAGAAATATCACTTTAATTAAGGT 865
Db

RESULT 14
AAT43259
ID AAT43259 standard; cDNA to mRNA; 2028 BP.
AC
XX AAT43259;
XX
DT 25-FEB-1997 (first entry)
XX
DE Type I macrophage scavenger receptor gene.
XX
KW Type I macrophage scavenger receptor; SRG; kidney; 293 cell;
XX cell substrate; cell attachment; atherosclerosis; ds.
XX Homo sapiens.
XX
FH Location/Qualifiers
FT 47..1402
FT /*tag= a
XX
XX WO9638725-A1.
XX
XX 05-DEC-1996.
XX
XX 30-MAY-1996; 96WO-US008081.
XX
XX 30-MAY-1995; 95US-00453117.
XX
XX (SMIK ) SMITHKLINE BEECHAM CORP.
XX
XX Lyeko PG, Elshourbagy NAE, Brawner ME;
XX
XX WPI; 1997-034505/03.
XX
XX P-PSDB; AAW08077.
XX
XX New modified human embryonic kidney 293 cells - transfected with a
XX mammalian scavenger receptor gene to enhance ability to attach to a solid
XX support.
XX
XX Example 1; Page 10-13; 32pp; English.
XX
XX Human type I and type II macrophage scavenger receptor genes (AAT43259
XX and AAT43260, respectively) can be used to transfect human embryonic
XX kidney 293 cells. The transfected cells show an enhanced ability to
XX attach to a solid support and are useful for screening cpds. for
XX biological activity, or for identifying antagonists of the scavenger
XX receptor gene, e.g. to develop agents for treatment of atherosclerosis
XX
XX Sequence 2028 BP; 653 A; 369 C; 434 G; 572 T; 0 U; 0 Other;
XX

Alignment Scores:
Pred. No.: 5.14e-85 Length: 2028
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 2 Gaps: 1

US-10-618-570-2 (1-400) x AAT43259 (1-2028)

Qy 1 MetalaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
|||||.....|||||.....|||||.....|||||.....|||||.....
Db 47 ATGGAGCAGTGGGATCACTTTTCAATCAACAGGAGGACACTGATAGCTGCTCCGAATCT 106
|||||.....|||||.....|||||.....|||||.....|||||.....
Qy 21 VallysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
|||||.....|||||.....|||||.....|||||.....|||||.....
Db 107 GTGAAATTTTGATGCTCGTCAATGACAGCTTTGCTTCTCCGAATCTCTAAACACAGCCCT 166
|||||.....|||||.....|||||.....|||||.....|||||.....

```

```

Qy 41 ThrLeuGlnGluArgMetLysSerTyrIysThrAlaLeuIleThrLeuTyrLeuIleVal 60
|||||.....|||||.....|||||.....|||||.....|||||.....
Db 167 TCCCTTCAAGAGAACTGAAGTCTCTCAAAAGCTGCACCTGATTGGCTTTTACCTCTCGTG 226
|||||.....|||||.....|||||.....|||||.....|||||.....
Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTrpGluThr 80
|||||.....|||||.....|||||.....|||||.....|||||.....
Db 227 TTTGCAGTTCTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTCGAAGTGGGAACG 286
|||||.....|||||.....|||||.....|||||.....|||||.....
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
|||||.....|||||.....|||||.....|||||.....|||||.....
Db 287 AAGAATTGCTCAGTTAGTTCAACTAATGCAAAATGATATACTCAAAAGTCTCAGCGGAAAA 346
|||||.....|||||.....|||||.....|||||.....|||||.....
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
|||||.....|||||.....|||||.....|||||.....|||||.....
Db 347 GGAATGACACGGAAGGAAATGAGATTTCAAGAAGTCTTTATGGAAACACATGAGCAAC 406
|||||.....|||||.....|||||.....|||||.....|||||.....
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
|||||.....|||||.....|||||.....|||||.....|||||.....
Db 407 ATGGAGAGAGATCCAGCATATTTTAGACATGGAAGCCAACTCATGGACACAGAGCAT 466
|||||.....|||||.....|||||.....|||||.....|||||.....
Qy 140 PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
|||||.....|||||.....|||||.....|||||.....|||||.....
Db 467 TTCCAAATTTTCAGCATGACAACTGATCAAAAGATTTAATGACATCTTCTGCAGCTAAGT 526
|||||.....|||||.....|||||.....|||||.....|||||.....
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
|||||.....|||||.....|||||.....|||||.....|||||.....
Db 527 ACCTTGTTTCTCAGTCCAGGACATGGGAATGCAATAGATGAAATCTCCAAGTCTCTTA 586
|||||.....|||||.....|||||.....|||||.....|||||.....
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
|||||.....|||||.....|||||.....|||||.....|||||.....
Db 587 ATAAGTTTGAATACCACTTGTGATTTGAGTCAACATAGAAATCTCAATGGCAAA 646
|||||.....|||||.....|||||.....|||||.....|||||.....
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
|||||.....|||||.....|||||.....|||||.....|||||.....
Db 647 ATCCAAGAGATACCTTCAACACACAGAGGAATCAGTAATTTAGAGGCGCTGTTTAC 706
|||||.....|||||.....|||||.....|||||.....|||||.....
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
|||||.....|||||.....|||||.....|||||.....|||||.....
Db 707 AATGTATCAGCAGAAATTTATGGCTATGAAAGAGAACAAAGTGCAATTTGGAAACAGGAATA 766
|||||.....|||||.....|||||.....|||||.....|||||.....
Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
|||||.....|||||.....|||||.....|||||.....|||||.....
Db 767 AAAGGAGAGTGAAGTACTGAAATACATCACTAATGATCTCAGACTGAAAGATTGGGAA 826
|||||.....|||||.....|||||.....|||||.....|||||.....
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
|||||.....|||||.....|||||.....|||||.....|||||.....
Db 827 CATTTCTCAGAGCTTGAGAAATATCACTTTAATTAAGGT 865
|||||.....|||||.....|||||.....|||||.....|||||.....

```

RESULT 15

AAF29164

ID AAF29164 standard; DNA; 2028 BP.

XX

AC AAF29164;

XX

DT 05-APR-2001 (first entry)

XX

DE Human DNA encoding type I scavenger receptor.

XX

KW Human; type I scavenger receptor; antibody; macrophage; ds.

XX

OS Homo sapiens.

XX

XX JP2000312595-A.

PN

XX 14-NOV-2000.

PD

PF 16-AUG-1991; 2000JP-00114233.

XX

PR 27-AUG-1990; 90JP-00222398.

XX

PR 16-AUG-1991; 91JP-00229728.

XX

PA (CHUS) CHUGAI PHARM CO LTD.

XX WPI; 2001-161981/17.
 DR P-PSDB; AAB49715.
 XX New anti-human scavenger receptor antibody useful for the identification
 PT and the determination of macrophages.
 XX Example 1; Page 5-7; 13pp; Japanese.
 XX This invention relates to a poly or monoclonal antibody which targets a
 CC human type I or type II scavenger receptor epitope. The anti-scavenger
 CC receptor antibody can be used for the identification and determination of
 CC macrophages. The present sequence represents DNA encoding the human type
 CC I scavenger receptor
 XX
 SQ Sequence 2028 BP; 653 A; 369 C; 434 G; 572 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 5,14e-85 Length: 2028
 Score: 1042.50 Matches: 197
 Percent Similarity: 97.2% Conservative: 41
 Best Local Similarity: 72.2% Mismatches: 34
 Query Match: 50.6% Indels: 1
 DB: 4 Gaps: 1

US-10-618-570-2 (1-400) x AAF29164 (1-2028)

Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
 Db 47 ATGGAGCAGTGGGATCACTTTACAAATCAACAGGAGGACACTGATAGCTGCTCGGAATCT 106
 Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
 Db 107 GTGAATTTGATGCTCGCTCAATGACAGCTTTGCTCTCGAATCTTAAACACAGCCCT 166
 Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuIleVal 60
 Db 167 TCCCTTCAAGAGAACTAGTCTCTTCAAGCTGACCTGATGCCCCTTACTCTCGTG 226
 Qy 61 PheValValLeuValProIleGlyIleValAlaAlaGlnLeuLeuLysTrpGluThr 80
 Db 227 TTTCAGTTCATCCCTCTCATTTGGAATAGTGCAGCTCAACTCTGAACTGAGGGAACG 286
 Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
 Db 287 AAGAAATGCTCAGTTAGTTCACTAATGCAATGATATATACTCAAGTCTCACGGGAAA 346
 Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
 Db 347 GGAATGACAGGAGAGGAATGAGATTTCAAGAGTCTTTATGGAACACATGAGCAAC 406
 Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
 Db 407 ATGGAGAAGAAATCCAGCATATTTAGACATGGAAGCCAACTCATGGACACAGAGCAT 466
 Qy 140 PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
 Db 467 TTCCAAAATTTCCAGCATGACCACTCATCAAGATTTAATGACATTTCTTCGACGCTAAGT 526
 Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleLeuGlyAspIleSerLysSerLeu 179
 Db 527 ACCTTGTTTCTCAGTCAGGACATGGGAATGCAATAGATGAATCTCCAGTCTCTTA 586
 Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
 Db 587 ATAAGTTTGAATACCACATCTGCTTATTTGAGCTCAACATAGAAAATCTGAATGGCAA 646
 Qy 200 ValGlnGluAsnAlaPheLysGlnGluMetArgLysLeuGluGluArgIleTyr 219
 Db 647 ATCCAAGAGAATACCTTCAACACAAAGAGAAATCAGTAAATTAGAGGCGGTTCAC 706
 Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnIle 239

Db 707 AATGTATCAGCAGAAATTTATGGCTATGAAAGAGACAAAGTGCATTTTGGAAACAGGAAATA 766
 Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
 Db 767 AAAGGAGAGTGAAGTACTGAAATACATCACTAATGATCTCAGACTGAAAAGATTGGGAA 826
 Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
 Db 827 CATTTCTAGACCTTGAGAAATATCACTTTAATTCAAGGT 865

Search completed: February 27, 2006, 09:40:04
 Job time : 659 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 27, 2006, 09:24:14 ; Search time 1029 Seconds
(without alignments)
3214.530 Million cell updates/sec

Title: US-10-618-570-2

Perfect score: 2059

Sequence: 1 MAQWDFPDQEDTDSCTES.....DWKATRVGINFTLRTOKE 400

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=abss/ABSSWEB spool/US10618570/runat 27022006 063901 13209/app query.fasta_1
-DB=Published Applications NA Main -QFWT=fastap -SUPFIX=rnpsbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs04
-USER=US10618570.ecgn 1 1026 @runat 27022006 063901 13209 -NCPU=6 -ICPU=3
-NO MMAP -NEG_SCORES=0 -WAIT -DSFBLCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2059	100.0	5177	8	US-10-618-570-1
2	1042.5	50.6	1077	6	US-10-326-186-5
3	1042.5	50.6	1367	6	US-10-305-720-1260
4	1042.5	50.6	2028	7	US-10-426-262-1
5	1042.5	50.6	2877	8	US-10-741-600-570
6	1042.5	50.6	3336	8	US-10-741-600-568
7	1042.5	50.6	3525	8	US-10-741-600-567

8	1042.5	50.6	3704	6	US-10-240-965-123	Sequence 123, Appl
9	1042.5	50.6	3719	8	US-10-741-600-569	Sequence 569, Appl
10	1042.5	50.6	3791	8	US-10-741-600-566	Sequence 566, Appl
11	994.5	48.3	1330	6	US-10-393-892-41	Sequence 41, Appl
12	994.5	48.3	1330	6	US-10-394-382-41	Sequence 41, Appl
13	688	33.4	459	7	US-10-733-031-1	Sequence 1, Appl
14	678	32.9	486	10	US-11-093-776-6	Sequence 6, Appl
15	678	32.9	576	10	US-11-093-776-2	Sequence 2, Appl
16	678	32.9	604	8	US-10-724-527-3	Sequence 3, Appl
17	670	32.5	540	9	US-10-893-576-57	Sequence 57, Appl
18	670	32.5	540	9	US-10-893-576-58	Sequence 58, Appl
19	495.5	24.1	119036	8	US-10-741-600-17721	Sequence 17721, A
20	282	13.7	201	8	US-10-741-600-13815	Sequence 13815, A
21	282	13.7	201	8	US-10-741-600-13824	Sequence 13824, A
22	282	13.7	201	8	US-10-741-600-13833	Sequence 13833, A
23	282	13.7	201	8	US-10-741-600-13844	Sequence 13844, A
24	282	13.7	201	8	US-10-741-600-13854	Sequence 13854, A
25	254	12.3	201	8	US-10-741-600-13814	Sequence 13814, A
26	254	12.3	201	8	US-10-741-600-13823	Sequence 13823, A
27	254	12.3	201	8	US-10-741-600-13832	Sequence 13832, A
28	254	12.3	201	8	US-10-741-600-13842	Sequence 13842, A
29	254	12.3	201	8	US-10-741-600-13850	Sequence 13850, A
30	204	9.9	1485	7	US-10-468-335-1	Sequence 1, Appl
31	204	9.9	1581	7	US-10-138-588-65	Sequence 65, Appl
32	204	9.9	3644	9	US-10-468-335-3	Sequence 3, Appl
33	201	9.8	2885	9	US-10-496-905-443	Sequence 443, Appl
34	178	8.6	201	8	US-10-741-600-48346	Sequence 48346, A
35	171.5	8.3	3900	7	US-10-311-623-22	Sequence 22, Appl
36	169	8.2	743	6	US-10-326-186-7	Sequence 7, Appl
37	168.5	8.2	713	6	US-10-326-186-9	Sequence 9, Appl
38	167	8.1	1176	9	US-10-671-995A-5	Sequence 5, Appl
39	166.5	8.1	1395	6	US-10-312-245-5	Sequence 5, Appl
40	155	7.5	201	8	US-10-741-600-13811	Sequence 13811, A
41	155	7.5	201	8	US-10-741-600-13820	Sequence 13820, A
42	155	7.5	201	8	US-10-741-600-13829	Sequence 13829, A
43	155	7.5	201	8	US-10-741-600-13838	Sequence 13838, A
44	155	7.5	201	8	US-10-741-600-13849	Sequence 13849, A
45	155	7.5	840	9	US-10-471-422-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-10-618-570-1
; Sequence 1, Application US/10618570
; Publication No. US20040185059A1
; GENERAL INFORMATION:
; APPLICANT: Yla-Herttuala, Seppo
; APPLICANT: Kulomaa, Markku
; APPLICANT: Lehtolainen, Paulina
; APPLICANT: Marjomaki, Varpu
; APPLICANT: Airenne, Kari
; TITLE OF INVENTION: Biotin-Binding Receptor Molecules
; FILE REFERENCE: GJE-48
; CURRENT APPLICATION NUMBER: US/10/618,570
; CURRENT FILING DATE: 2003-07-11
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US/09/622,804
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent version 3.1
; SEQ ID NO 1
; LENGTH: 5177
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant cDNA encoding fusion protein
; NAME/KEY: CDS
; LOCATION: (1071)..(2270)
; OTHER INFORMATION:
US-10-618-570-1
Alignment Scores:

```
Pred. No.: 1.34e-194 Length: 5177
Score: 2059.00 Matches: 400
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-618-570-2 (1-400) x US-10-618-570-1 (1-5177)

Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 1071 ATGGCACAGTGGGATGACTTTCTCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCT 1130

Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisLeuProLysAsnGlyPro 40
Db 1131 GTGAAGTTCGATCTCGCTCAGTGACGCTTGCTTCTCCCATCTCTAAATAATGGCCCA 1190

Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60
Db 1191 ACTCTTCAAGAGAGGATGAAGTCTTATAAACTGCACCTGATCACCTTTATCTCATTTGTG 1250

Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTrpGluThr 80
Db 1251 TTTGTAGTTCGTGCTCCCATCATTTGGCATAGTGGCAGCTCAGCTCTCTGAAATGGGAACG 1310

Qy 81 LysAsnCysThrValGlySerValAsnAlaAspLysSerProSerProGluGlyLysGly 100
Db 1311 AAGAAATTCACGGTGGCTCAGTTANTGCAGATATATCTCCAAGTCCGGAAGGCCAAGGA 1370

Qy 101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120
Db 1371 AATGSCAGTGAAGATGAATGAGATTTTCGAGAAGCTGTGATGGAACGCATCAGCAACATG 1430

Qy 121 GluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPhe 140
Db 1431 GAAAGCAGAAATCCAGTATCTTTTCAGATAATGAAGCCAAATCTCCTAGATGCTAAGAAATTC 1490

Qy 141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
Db 1491 CAAATTTTCAGATAAACAACTGATCAAGATTTAATGATGTTCTTTTCCAGCTAAATTC 1550

Qy 161 LeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeuVal 180
Db 1551 TTACTTTCTCCATCCAGGACATCAGAAATATCATAGGGGATATCTCCAAGTCAATAGTA 1610

Qy 181 GlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArgVal 200
Db 1611 GGTCTGAACACCCACAGTACTTGAATTTGCAGTTTCAGTATTGAAACACTGAATGGCAGAGTC 1670

Qy 201 GlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyrAsn 220
Db 1671 CAAGAGATGCATTTTAAACAACAGAGAGATGCGTAATTAGAGGCGGTATATACAAAT 1730

Qy 221 AlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIleLys 240
Db 1731 GCATCAGCAGAAATTAAGTCTCTAGATGAAAAACAAGTATATTTGGACAGGAAATRAAA 1790

Qy 241 GlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGluHis 260
Db 1791 GGGGAAATGAACACTTGAATAATATCATCAATGATCTGAGGCTGAAGGATGGGAAACAT 1850

Qy 261 SerGlnThrLeuLysAsnIleThrLeuLeuGlnGlyAlaArgLysCysSerLeuThrGly 280
Db 1851 TCTCAGACATTTGAAAAATATACATTTACTCCAAGGTGCCAGAAAGTGTCTGCTGACTGGG 1910

Qy 281 LysTrpThrAsnAspLeuGlySerAsnMetThrIleGlyAlaValAsnSerArgGlyGlu 300
Db 1911 AAATGACCAACGATCTGGGCTCCAAACATGACCATCGGGGCTGTGAAACAGCAGAGGTGA 1970

Qy 301 PheThrGlyThrTyrIleThrAlaValThrAlaThrSerAsnGluIleLysGluSerPro 320
Db 1971 TTCACAGCCACTCATCACAGCCGTAACAGCCACATCAAAATGAGATCAAAAGAGTCACCA 2030
```

```
Qy 321 LeuHisGlyThrGlnAsnThrIleAsnLysArgThrGlnProThrPheGlyPheThrVal 340
Db 2031 CTGCATGGGACACAAAACACCATCAACAGAGGACCCAGCCACCTTTGGCTTCAACCGTC 2090

Qy 341 AsnTrpLysPheSerGluSerThrValPheThrGlyGlnCysPheIleAspArgAsn 360
Db 2091 AATTGGAAGTTTTCAGAGTCCACCACCTGCTTTCACGGGCCAGTGTCTTCATAGACAGGAAT 2150

Qy 361 GlyLysGluValLeuLysThrMetTrpLeuLeuArgSerValAsnAspIleGlyAsp 380
Db 2151 GGGAAAGAGGTCCTGAAGACCATGTGCTGCTGGGTCAAGTGTAAATGACATTTGGTGAT 2210

Qy 381 AspTrpLysAlaThrArgValGlyIleAsnIlePheThrArgLeuArgThrGlnLysGlu 400
Db 2211 GACTGGAAAGCTACCAGGTGCGCATCAACATCTTCACTCGCCTCGGCACACAGAGGAG 2270

RESULT 2
US-10-326-186-5
; Sequence 5, Application US/10326186
; Publication No. US20030119149A1
; GENERAL INFORMATION:
; APPLICANT: Reddy, Pranhitha
; TITLE OF INVENTION: Trimeric Recombinant Polypeptides
; FILE REFERENCE: 3398-A
; CURRENT APPLICATION NUMBER: US/10/326,186
; CURRENT FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1077
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1077)
; OTHER INFORMATION:
US-10-326-186-5

Alignment Scores:
Pred. No.: 6.01e-94 Length: 1077
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservatives: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 6 Gaps: 1

US-10-618-570-2 (1-400) x US-10-326-186-5 (1-1077)

Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 1 ATGGAGCAGTGGGATCACTTTCAATCAACAGAGGACACTGATAGTGTCCGNAATCT 60

Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisLeuProLysAsnGlyPro 40
Db 61 GTGAAATTTGATGCTCGCTCAATGACAGCTTGTCTCTCGAAATCTCTAAACACAGCCCT 120

Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60
Db 121 TCCCTTCAAGAGAACTGGAAGTCTCTCAAAAGCTGCATGATTCCTCTTCTCTCTCTG 180

Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTrpGluThr 80
Db 181 TTTGCAGTTCTCATCCCTCTCATTTGGATAGTGGCAGCTCAACTCTCTGAAAGTGGGAACG 240

Qy 81 LysAsnCysThrValGlySerValAsnAla---AspLysSerProSerProGluGlyLys 99
Db 241 AAGAATTTGCTCAGTTAGTTCAACTTAATGCAAAATGATATACTCAAAAGTCTCACGGGAAA 300

Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 301 GGAATCAGACGGAAGGAAATGAGATTTCAAGAAAGTCTTTATGGAAACATGAGCAAC 360

Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
```

```

Db 361 ATGGAGAGAGAGATCCAGCATATTTTACATGGAAGGCAACCTCATGGACACAGAGCAT 420
Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 421 TTCCAAATTTTCAGCATGACAACTGATCAAGATTTTAATGACATTTCTTCGACGTAAGT 480
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerIleSerLeu 179
Db 481 ACCTTGTTTCTCAGTCAGGAGACATGGGATGCAATAGATGAATCTCCAAGTCCCTTA 540
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 541 ATAGTTTGAATACCACTATGCTTGAATGCTCAACATAGAAAATCTGAATGGCAAA 600
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 601 ATCCAAAGAGATPACCTTCAACAACAGAGGAATCAGTAAATTAGAGGAGCGGTGTATAC 660
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
Db 661 AATGATATCAGCAGAAATTTATGCTATGAAGAAGAACAAAGTGCATTTTGGAAACAGGAATA 720
Qy 240 LysGlyGluMetLysLeuLeuAsnIleThrAsnAspLeuAtcLeuLysAspTrpGlu 259
Db 721 AAAGAGAGAGTGAAGTACTGAATAACATCACTAATGATCTCAGACTGAAAGATTGGAA 780
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 781 CATCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 819

```

RESULT 3

```

US-10-305-720-1260
; Sequence 1260, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; PRIOR FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 1260
; LENGTH: 1367
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank ID No. US20040010136A1 g219991
US-10-305-720-1260

```

Alignment Scores:

```

Pred. No.: 8,46e-94 Length: 1367
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 6 Gaps: 1

```

US-10-618-570-2 (1-400) x US-10-305-720-1260 (1-1367)

```

Qy 1 MetAlaGlnTrpAspAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 67 ATGGAGCAGTGGATTCATTTTCAATCAACAGAGGAGCACTGATAGTCTCCGAATCT 126
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 127 GTGAATTTGATGCTCGCTCATGACAGCTTGCTTCTCCGATCTCCGATTTAAACAGCCCT 186
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60

```

```

Db 187 TCCCTTCAAGAGAACTGAAGTCTCTCAAGAGCTGCACCTGATTCCTCTCTCTCTG 246
Qy 61 PheValValLeuValProIleIleGlyIleValAlaalaGlnLeuLeuLysTrpGluThr 80
Db 247 TTTGACGTTTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTCTGAAGTGGGAACG 306
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 307 AAGAATTTGCTCAGTTAGTTCAACTAATGCAATATGATATACTCAAGTCTCTCACGGGAANA 366
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 367 GGAATGACACGAGGAGGAATGAGATTTCAAGAGTCTTTATGGAACACATGAGCAAC 426
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 427 ATGGAGAAGAGATCCAGCATATTTTAGACATGGAAGCCCACTCATGCACACAGAGCAT 486
Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 487 TTCCAAATTTTCAGCATGACAACTGATCAAGATTTTAATGACATTTCTCTCGACGTAAGT 546
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerIleSerLeu 179
Db 547 ACCTTGTTTCTCAGTCCAGGAGCATGGGAATGCAATAGATGAATCTCCAAGTCCCTTA 606
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 607 ATAAGTTTGAATACCACTGCTTGAATTTGACGCTCAACATAGAAAATCTGAATGGCAAA 666
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 667 ATCCAAGAGATACCTTCAACAACAAGAGAAATCAGTAAATTAGAGGAGCGGTGTATAC 726
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
Db 727 AATGATATCAGCAGAAATTTATGCTATGAAGAAGAACAAAGTGCATTTTGGAAACAGGAATA 786
Qy 240 LysGlyGluMetLysLeuLeuAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
Db 787 AAAGAGAGAGTGAAGTACTGATGAATACATCACTAATGATCTCAGACTGAAAGATTGGAA 846
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 847 CATCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 885

```

RESULT 4

```

US-10-426-262-1
; Sequence 1, Application US/10426262
; Publication No. US20040018521A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jianfeng
; APPLICANT: Meyers, Deborah
; APPLICANT: Zheng, Sigun
; APPLICANT: Walsh, Patrick C.
; APPLICANT: Isaacs, William B.
; APPLICANT: Bleeker, Eugene
; TITLE OF INVENTION: MUTATIONS IN THE MACROPHAGE SCAVENGER RECEPTOR 1 GENE ALTER RISK
; TITLE OF INVENTION: OF PROSTATE CANCER, ASTHMA, AND CARDIOVASCULAR DISEASE
; FILE REFERENCE: 9151-23
; CURRENT APPLICATION NUMBER: US/10/426,262
; CURRENT FILING DATE: 2003-05-01
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2028
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1402)
US-10-426-262-1

```

```
Alignment Scores: 1.49e-93 Length: 2028
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 7 Gaps: 1

US-10-618-570-2 (1-400) x US-10-426-262-1 (1-2028)

Qy 1 MetAlaGlnTrpAspAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
Db 47 ATGGAGCAGTGGGATCACTTTCCACAATCAACAGAGGACACTGATAGCTGCTCCGAATCT 106

Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 107 GTGAAATTTGATGCTCGCTCAATGACAGCTTTGCTTCCTCCGAATCTCTTAAACACAGCCCT 166

Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuLeuLeuVal 60
Db 167 TCCCTTCAAGAGAACTGAGTCTTCAAGCTGCACTGATGCTTACCTCTCTCGTG 226

Qy 61 PheValValLeuValProIleGlyIleValAlaAlaGlnLeuLeuLysTrpGluThr 80
Db 227 TTTCAGTCTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTTGAAGTGGGAAACG 286

Qy 81 LysAsnCysThrValGlySerValAsnAla---AspLysSerProSerProGluGlyLys 99
Db 287 AAGAAATTCAGTTAGTTAGTCAACTAATGCAATGATATAACTTCAAGTCTCACGGGAAA 346

Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 347 GGAATGACAGCGAAGAGAAATGAGATTTCAAGAGTCTTTATGGAACACATGAGCAAC 406

Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGlnAlaAsnLeuLeuAspAlaLysAsn 139
Db 407 ATGGAGAGAGATCCAGCATATTTAGACATGGAGGCAACCTCATGGACACAGAGCAT 466

Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 467 TTCCAAAATTTTCAGCATGACAACTGATCAAGAGATTTAATGACATTTCTTCGAGCTAAGT 526

Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
Db 527 ACCTTGTTTCTCAGTCAGGACATGGGAATGCAATAGATGAAATCTCCAAGTCTCTTA 586

Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 587 ATAAGTTTGAATACCATCTGTTGATTTGAGCTCAACATAGAAAATCTGAATGGCAAA 646

Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 647 ATCCAAGAGATATACCTTCAAAACAACAGAGGAAATCAGTAAATTTAGAGGCGGTGTTTAC 706

Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGluIle 239
Db 707 AATGATCAGCAGAAATTTATGGCTATGAAAGAGAAACAAAGTGCAATTTGGAACAGGAATA 766

Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
Db 767 AAAGAGAGATGGAAGTACTGAATACATCACTAATGATCTCAGACTGAAGATTTGGAA 826

Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 827 CATTCACAGACCTTGAGAAATATACATCTTAATTTCAAGT 865

RESULT 5
US-10-741-600-570
; Sequence 570, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
```

```
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10741.600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 570
; LENGTH: 2877
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-570

Alignment Scores: 2.46e-93 Length: 2877
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 8 Gaps: 1

US-10-618-570-2 (1-400) x US-10-741-600-570 (1-2877)

Qy 1 MetAlaGlnTrpAspAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
Db 175 ATGGAGCAGTGGGATCACTTTCCACAATCAACAGAGGACACTGATAGCTGCTCCGAATCT 234

Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 235 GTGAAATTTGATGCTCGCTCAATGACAGCTTTGCTTCCTCCGAATCTCTTAAACACAGCCCT 294

Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuLeuLeuVal 60
Db 295 TCCCTTCAAGAGAACTGAGTCTTCAAGCTGCACTGATGCTTACCTCTCTCGTG 354

Qy 61 PheValValLeuValProIleGlyIleValAlaAlaGlnLeuLeuLysTrpGluThr 80
Db 355 TTTCAGTCTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTTGAAGTGGGAAACG 414

Qy 81 LysAsnCysThrValGlySerValAsnAla---AspLysSerProSerProGluGlyLys 99
Db 415 AAGAAATTCAGTTAGTTAGTCAACTAATGCAATGATATAACTTCAAGTCTCACGGGAAA 474

Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 475 GGAATGACAGCGAAGAGAAATGAGATTTCAAGAGTCTTTATGGAACACATGAGCAAC 534

Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGlnAlaAsnLeuLeuAspAlaLysAsn 139
Db 535 ATGGAGAGAGATCCAGCATATTTAGACATGGAGGCAACCTCATGGACACAGAGCAT 594

Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 595 TTCCAAAATTTTCAGCATGACAACTGATCAAGAGATTTAATGACATTTCTTCGAGCTAAGT 654

Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
Db 655 ACCTTGTTTCTCAGTCCAGGACATGGGAATGCAATAGATGAAATCTCCAAGTCTCTTA 714

Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 715 ATAAGTTTGAATACCATCTGTTGATTTGAGCTCAACATAGAAAATCTGAATGGCAAA 774

Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 775 ATCCAAGAGATATACCTTCAAAACAACAGAGGAAATCAGTAAATTTAGAGGCGGTGTTTAC 834

Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGluIle 239
Db 835 AATGATCAGCAGAAATTTATGGCTATGAAAGAGAAACAAAGTGCAATTTGGAACAGGAATA 894

Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
```

```
Db 895 AAAGGAGAGTGAAGTACTGAATACATCACTAATGATCTCAGACTGAAGATTGGAA 954
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 955 CATTCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 993

RESULT 6.
US-10-741-600-568
; Sequence 568, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 568
; LENGTH: 3336
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-568

Alignment Scores:
Pred. No.: 3,04e-93 Length: 3336
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 8 Gaps: 1

US-10-618-570-2 (1-400) x US-10-741-600-568 (1-3336)
Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 160 ATGGAGCAGTGGGATCACTTTCAATCAATCAACAGGAGGACACTGATAGCTGCCGAATCT 219
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 220 GTGAATTTGATGCTCGTCAATGACAGCTTTGCTTCTCGAATCTTAAACAGCCCT 279
Qy 41 ThrLeuGlnGluArgMetLysSerTyrlsThrAlaLeuIleThrLeuTyLeuVal 60
Db 280 TCCCTTCAAGAGAACTCAAGTCTTCAAGCTGCACTGATTCCTTACCTTCTCTG 339
Qy 61 PheValValLeuValProIleGlyIleValAlaAlaGlnLeuLeuTyTrpGluThr 80
Db 340 TTTCAGGTTCTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCCTGAAGTGGGAACG 399
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 400 AAGAAATGCTCAGTTAGTTCACTAATGCAATGATATGATTAATCAAGTCTCACGGGAAA 459
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 460 GGAAATGACGGAAGAGAAATGAGATTTCAAGAAGTCTTTATGGAACACATCAGCAAC 519
Qy 120 MetGluSerArgIleGlnTyLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 520 ATGGAGAAGAAATCCAGCATATTTTAGCATGGAAGCAACCTCATGGACACAGAGCAT 579
Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 580 TTCCAAATTTTCAGTCTGACAACTGATCAAGATTTAATGATTCCTTCTCAGCTAAGT 639
Qy 160 SerLeuLeuSerSerIleGlnGlnHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
Db 640 ACCTTGTCTTCTCAGTCCAGGACATGGGAATGCAATAGATGAATCTTCAAGTCTCTTA 699
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 955 CATTCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 993
```

```
Db 700 ATAAGTTTGAATACCACTTGTGTTGATTGCGCTCAACATAGAAAATCTGAATGGCAA 759
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTy 219
Db 760 ATCCAGAGAAATACCTTTCAACAAACAGAGGAAATCAGTAATATTAGAGGAGCGTGTTC 819
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyLeuGluGlnGluIle 239
Db 820 AATGTATRCGAGAAATTTATGGCTATGAAAGAACAAACAGTGCATTTTGGAAACAGGAATA 879
Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
Db 880 AAGAGAGAGTGAAGTACTGATTAACATCACTAATGATCTCAGACTGAAAGATTGGAA 939
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 940 CATTCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 978

RESULT 7
US-10-741-600-567
; Sequence 567, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 567
; LENGTH: 3525
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-567

Alignment Scores:
Pred. No.: 3,29e-93 Length: 3525
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 8 Gaps: 1

US-10-618-570-2 (1-400) x US-10-741-600-567 (1-3525)
Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 160 ATGGAGCAGTGGGATCACTTTCAATCAATCAACAGGAGGACACTGATAGCTGCCGAATCT 219
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 220 GTGAATTTGATGCTCGTCAATGACAGCTTTGCTTCTCGAATCTTAAACAGCCCT 279
Qy 41 ThrLeuGlnGluArgMetLysSerTyrlsThrAlaLeuIleThrLeuTyLeuVal 60
Db 280 TCCCTTCAAGAGAACTCAAGTCTTCAAGCTGCACTGATTCCTTACCTTCTCTG 339
Qy 61 PheValValLeuValProIleGlyIleValAlaAlaGlnLeuLeuTyTrpGluThr 80
Db 340 TTTCAGGTTCTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCCTGAAGTGGGAACG 399
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 400 AAGAAATGCTCAGTTAGTTCACTAATGCAATGATATGATTAATCAAGTCTCACGGGAAA 459
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 460 GGAAATGACGGAAGAGAAATGAGATTTCAAGAAGTCTTTATGGAACACATCAGCAAC 519
Qy 120 MetGluSerArgIleGlnTyLeuSerAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 139
Db 955 CATTCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 978
```

```

Db 520 ATGGAGAGAGAGATCCAGCATATTTTACATGGAAGCCAACTCATGGACACAGAGCAT 579
Qy 140 PheGlnAenPheSerIleThrAspGlnArgPheAenAspValLeuPheGlnLeuAen 159
Db 580 TTCCAAAATTTTCAGCATGACAACTGATCAAAAGATTTAATGACATTTCTTCGACGCTAAGT 639
Qy 160 SerLeuLeuSerSerIleGlnGlnHisGlnAenIleGlyAspIleSerLysSerLeu 179
Db 640 ACCITGTTTCTCAGTCAGGACATGGGAATGCAATAGATGAATCTCCAGTCCCTTA 699
Qy 180 ValGlyLeuAenThrValLeuAspLeuGlnPheSerIleGluThrLeuAenGlyArg 199
Db 700 ATAAGTTGAATACCACTGCTTGTGATTTGAGCTCAACATAGAAAATCTCAATGGCAAA 759
Qy 200 ValGlnGluAenAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 760 ATCCAAAGAGATACCTTCAACAAACAGAGAAATCAGTAATTAGAGGCGGTGTTTAC 819
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
Db 820 AATGTATCCGAGAAATTTATGCTATGAAGAGACAACTGATTTGGAACAGGAATA 879
Qy 240 LysGlyGluMetLysLeuAenAenIleThrAenAspLeuArgLeuLysAspTyrGlu 259
Db 880 AAAGGAGAGTGAAGTACTGAATAACATCACTAATGATCTCAGACTGAAAGATTGGAA 939
Qy 260 HisSerGlnThrLeuLysAenIleThrLeuLeuGlnGly 272
Db 940 CATTCACAGCCTTGAGAAATATCACTTTAATCAAGGT 978

```

RESULT 8

```

US-10-240-965-123
; Sequence 123, Application US/10240965
; Publication No. US20030165924A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: SHIFFMAN, Dov
; APPLICANT: SOMOGYI, Roland
; APPLICANT: LAWN, Richard M.
; APPLICANT: SEILHAWER, Jeffrey J.
; APPLICANT: PORTER, Gordon J.
; APPLICANT: MIKITA, Thomas
; APPLICANT: TAI, Julie
; TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
; FILE REFERENCE: PA-0025 PCT
; CURRENT APPLICATION NUMBER: US/10/240,965
; PRIORITY FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: 60/195,106
; PRIOR FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 276
; SOFTWARE: PERL Program
; SEQ ID NO 123
; LENGTH: 3704
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030165924A1 344240.2
; NAME/KEY: unsure
; LOCATION: 3522, 3634-3635, 3699
; OTHER INFORMATION: a, t, c, g, or other
US-10-240-965-123

```

Alignment Scores:

```

Pred. No.: 3,54e-93 Length: 3704
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 6 Gaps: 1

```

US-10-618-570-2 (1-400) x US-10-240-965-123 (1-3704)

```

Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
Db 55 ATGGAGCAGTGGGATCACTTTTCAATCAACAGAGGACACTGATAGCTGCTCGCATCT 114
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 115 GTGAAATTTGATGCTCGCTCAATGACAGCTTTGCTTCTCCGAACTCTTAAACACAGCCT 174
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60
Db 175 TCCCTTCAAGAGAACTGAAGTCTCTCAAAAGCTGCACCTGATTCCTTATCCCTCTCGTG 234
Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTyrGluThr 80
Db 235 TTTCAGCTTCTCATCCCTCTCATTTGGATAGTGGCAGCTCACTCTCGAAGTGGGAACG 294
Qy 81 LysAenCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 295 AAGAAATTTGCTCAGTTAGTTCAACTTAATGCAATATGATATACTCAAAAGTCTCACGGGAAA 354
Qy 100 GlyAenGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 355 GGAATATGACAGCGAAGAGGAATGAGATTTCAAGAACTCTTTATGGAAACACATGAGCAAC 414
Qy 120 MetGluSerArgIleGlnTyrLeuSerAsnGluAlaAenLeuLeuAspAlaLysAen 139
Db 415 ATGGAGAGAGATTCAGCATATTTTAGACATGGAAGCCAACTCATGGACACAGAGCAT 474
Qy 140 PheGlnAenPheSerIleThrAspGlnArgPheAenAspValLeuPheGlnLeuAen 159
Db 475 TTCCAAAATTTTCAGCATGACAACTGATCAAAAGATTTAATGACATTTCTTCGACGTAAGT 534
Qy 160 SerLeuLeuSerSerIleGlnGlnHisGluAenIleIleGlyAspIleSerLysSerLeu 179
Db 535 ACCTTGTGTTTCTCAGTCCAGGACATGGAATGCAATAGATGAATCTCCAAAGTCTCTTA 594
Qy 180 ValGlyLeuAenThrValLeuAspLeuGluPheSerIleGluThrLeuAenGlyArg 199
Db 595 ATAAGTTTGAATACCACATTTGCTTATTCAGCTCAACATAGAAAATCTGAATGGCAAA 654
Qy 200 ValGlnGluAenAlaPheLysGlnGlnGluMetArgLysLeuGluArgIleTyr 219
Db 655 ATCCAGAGAGATACCTTCAAAACAAAGAGAGAAATCAGTAAATTTAGAGGAGCGGTGTTTAC 714
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
Db 715 AATGTATCAGCAGAAATTTATGCTATGAAAGAGAAACAAAGTGCATTTGGAAACAGGAATA 774
Qy 240 LysGlyGluMetLysLeuLeuAenAenIleThrAenAspLeuArgLeuLysAspTyrGlu 259
Db 775 AAAGGAGAGTGAAGTACTGAATATACATCACTAATGATCTCAGACTGAAAGATTGGAA 834
Qy 260 HisSerGlnThrLeuLysAenIleThrLeuLeuGlnGly 272
Db 835 CATTCACAGCCTTGAGAAATATCACTTTAATCAAGGT 873

```

RESULT 9

```

US-10-741-600-569
; Sequence 569, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 569
; LENGTH: 3719
; TYPE: DNA

```


ORGANISM: Homo sapiens
US-10-741-600-569

Alignment Scores:
Pred. No.: 3,56e-93 Length: 3719
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservations: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 8 Gaps: 1

US-10-618-570-2 (1-400) x US-10-741-600-569 (1-3719)

```

Qy 1 MetalaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 175 ATGGAGCAGTGGGATCACTTTTCAATCAACAGGAGGACACTGATAGCTCTCGAATCT 234

Qy 21 VallysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 235 GTGAATTTGATGCTCGCTCAATGACCTTTGCTCTCCGAATCTCAAAACAGCCCT 294

Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
Db 295 TCCCTTCAAGAGAACTGAGTCTTCAAGCTGACCTGATTCCTTACTCTCTCGTG 354

Qy 61 PheValValLeuValProIleGlyLeValAlaLalaGlnLeuLeuTyrTrpGluThr 80
Db 355 TTTCAGTCTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTCAAGTGGGAAACG 414

Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 415 AAGAATTCCTCAGTTAGTTCAGTCACTTAATCAATGATATATCACTCAAGTCTCACGGGAAA 474

Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 475 GGAATGACAGGAGAGAAATGAGATTTCAAGAGTCTTTATGGAACACATGAGCAAC 534

Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 535 ATGGAGAAGAGAAATCCAGCATATTTAGACATGGAAGCAACCTCATGGACACAGAGCAT 594

Qy 140 PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 595 TTCCAAAATTTTCAGCATGACAACTGATCAAAAGATTTAATGACATTTCTTCGACCTAAGT 654

Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleGlyAspIleSerLysSerLeu 179
Db 655 ACCTTGTTCTCTAGTCAGGACATGGGAATGCAATAGATGAAATCTCAAGTCTCTTA 714

Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 715 ATAAGTTTGAATACCACATTCCTGATTTGACGCTCAACATAGAAAATCTCAATGGCAAA 774

Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 775 ATCCAAAGAGATACCTTCAAAACAAAGAGGAAATCAGTAAATTTAGAGGACGCTGTTTAC 834

Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
Db 835 AATGTATCRGAGAAATTTATGCTATGAAAGAGAAACAGTGCAATTTGGAACAGGAATA 894

Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
Db 895 AAAGAGAGAGTGAAGTACTGAATAACATCACTAATGATCTCAGACTGAAAGATTTGGAA 954

Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 955 CATTCAGACCTTGAGAAATATCACTTTAATTTCAAGGT 993

```

RESULT 10

US-10-741-600-566
; Sequence 566, Application US/10741600
; Publication No. US20050026169A1

GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001499
 ; CURRENT APPLICATION NUMBER: US/10/741,600
 ; CURRENT FILING DATE: 2003-12-22
 ; NUMBER OF SEQ ID NOS: 73997
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 566
 ; LENGTH: 3791
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-741-600-566

Alignment Scores:
 Pred. No.: 3,56e-93 Length: 3791
 Score: 1042.50 Matches: 197
 Percent Similarity: 87.2% Conservations: 41
 Best Local Similarity: 72.2% Mismatches: 34
 Query Match: 50.6% Indels: 1
 DB: 8 Gaps: 1

US-10-618-570-2 (1-400) x US-10-741-600-566 (1-3791)

```

Qy 1 MetalaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 160 ATGGAGCAGTGGGATCACTTTTCAATCAACAGGAGGACACTGATAGCTCTCGAATCT 219

Qy 21 VallysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 220 GTGAATTTGATGCTCGCTCAATGACAGCTTTGCTCTCCGAATCTCAAAACAGCCCT 279

Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuTyrLeuVal 60
Db 280 TCCCTTCAAGAGAACTGAGATTTTCAAGTCTCTCAAGTCTCACTGATTCCTTACTCTCTG 339

Qy 61 PheValValLeuValProIleGlyLeValAlaLalaGlnLeuLeuTyrTrpGluThr 80
Db 340 TTTCAGTCTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTCAAGTGGGAAACG 399

Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 400 AAGAATTTGCTCAGTTAGTTCAGTCACTTAATGCAATGATATATCACTCAAGTCTCACGGGAAA 459

Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 460 GGAATGACAGGAGAGAAATGAGATTTCAAGAGTCTTTATGGAACACATGAGCAAC 519

Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 520 ATGGAGAAGAGAAATCCAGCATATTTTAGACATGGAAGCAACCTCATGGACACAGAGCAT 579

Qy 140 PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 580 TTCCAAAATTTTCAGCATGACAACTGATCAAAAGATTTAATGACATTTCTTCGACCTAAGT 639

Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleGlyAspIleSerLysSerLeu 179
Db 640 ACCTTGTTCTCTAGTCCAGGACATGGGAATGCAATAGATGAAATCTCAAGTCTCTTA 699

Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 700 ATAAGTTTGAATACCACATTCCTGATTTGACGCTCAACATAGAAAATCTGAATGGCAAA 759

Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 760 ATCCAAAGAGATACCTTCAAAACAAAGAGGAAATCAGTAAATTTAGAGGACGCTGTTTAC 819

Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
Db 820 AATGTATCRGAGAAATTTATGCTATGAAAGAGAAACAGTGCAATTTGGAACAGGAATA 879

```

```
Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
Db 880 AAAGGAGAGTCAAGTACTGATTAACATCACTAATGATCTCAGACTGAAGATTGGAA 939

Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 940 CATCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 978

RESULT 11
US-10-393-892-41
; Sequence 41, Application US/10393892
; Publication No. US20030186302A1
; GENERAL INFORMATION:
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: COLORECTAL CANCER DIAGNOSTICS
; FILE REFERENCE: CDS 267 US NP
; CURRENT APPLICATION NUMBER: US/10/393,892
; PRIOR FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: 60/368,798
; PRIOR FILING DATE: 2002-03-29
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 41
; LENGTH: 1330
; TYPE: DNA
; ORGANISM: human
US-10-393-892-41

Alignment Scores:
Pred. No.: 5,12e-89 Length: 1330
Score: 994.50 Matches: 189
Percent Similarity: 85.3% Conservative: 44
Best Local Similarity: 69.2% Mismatches: 39
Query Match: 48.3% Indels: 1
DB: 6 Gaps: 1

US-10-618-570-2 (1-400) x US-10-393-892-41 (1-1330)
Qy 1 MetAlaGlnTrpAspAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
Db 1 ATGGCGCAGTGGGACAGCTTCACTGATCAACAGGAGGACACTGATAGCTGTTCAGAACTCT 60

Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 61 GTGAAGTTTGATGCTCGTCCAAACACAGCTTGTCTCCCAATCTTAAATATGGCCCT 120

Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60
Db 121 CCACCTTCAAGAAAGCTGAATCTTCAAGCTGCACCTGATTCGCCCTTTATCTCCTTGTG 180

Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTrpGluThr 80
Db 181 TTTGCTGTCTCATCTTCACTATCATGCAATATATGCAACAGTGTATCTCAGTCTCTCTGGGAATG 240

Qy 81 LysAsnCysThrValGlySerValAsnAlaAsp---IleSerProSerProGluGlyLys 99
Db 241 AAGAATTGCACAGTTGGTTGTTCAATTAATGCAACAGTGTATCTCAGTCTCTCTGGGAATG 300

Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 301 GGAAATGACAGTGAAGATGAAGTGAATTCGAGAAGTTGTTATGGAACACATTAGCAAG 360

Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 361 ATGGAGAAAAGAATCAATATATTTTCAGATCTGAAAGAAAATCTCGTAGATTCAGAGCAT 420

Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 421 TTTCAAAATTTTCAGTGCAACTGATCAACGATTTGCTGATGTTCTTCTCACTAAGT 480

Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
Db 481 ACCTTGGTTCCACAGTCCAGGCACATGGGAATGCGGTAGATGAATCACCAGGTCCTTTA 540
```

```
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 541 ATAACTGTGAATACACGCTGCTTGATTTGCACCTCTATGTAGAAACACTGAATGTCAA 600

Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 601 TTCCAGGAGAAATACACTTAAAGGGCAAGAGAAATCAGCAAAATTAAGGAGCGTGTGCAC 660

Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
Db 661 AATGCATCAGCAGAAATTAATGCTCTATGAAAGAGAACAAAGTGCATTTTGGAAACAGAAATA 720

Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
Db 721 AAAGAGAGTGAAGTCTGAAATTAACATCACTAATGATCTCAGGCTGAAAGATTGGAA 780

Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 781 CATCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 819
```

RESULT 12

```
US-10-394-382-41
; Sequence 41, Application US/10394382
; Publication No. US20030186303A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Yixin
; TITLE OF INVENTION: COLORECTAL CANCER DIAGNOSTICS
; FILE REFERENCE: CDS 266 US NP
; CURRENT APPLICATION NUMBER: US/10/394,382
; PRIOR FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: 60/368,687
; PRIOR FILING DATE: 2002-03-29
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 41
; LENGTH: 1330
; TYPE: DNA
; ORGANISM: human
US-10-394-382-41
```

```
Alignment Scores:
Pred. No.: 5,12e-89 Length: 1330
Score: 994.50 Matches: 189
Percent Similarity: 85.3% Conservative: 44
Best Local Similarity: 69.2% Mismatches: 39
Query Match: 48.3% Indels: 1
DB: 6 Gaps: 1
```

US-10-618-570-2 (1-400) x US-10-394-382-41 (1-1330)

```
Qy 1 MetAlaGlnTrpAspAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
Db 1 ATGGCGCAGTGGGACAGCTTCACTGATCAACAGGAGGACACTGATAGCTGTTCAGAACTCT 60

Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 61 GTGAAGTTTGATGCTCGCTCAATACAGCTTGTGTTCCCAATCTTAAATATGGCCCT 120

Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60
Db 121 CCACCTTCAAGAAAGCTGAATCTTCAAGCTGCACCTGATTCGCCCTTTATCTCCTTGTG 180

Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTrpGluThr 80
Db 181 TTTGCTGTCTCATCTTCACTATCATGCAATATATGCAACAGTGTATCTCAGTCTCTCTGGGAATG 240

Qy 81 LysAsnCysThrValGlySerValAsnAlaAsp---IleSerProSerProGluGlyLys 99
Db 241 AAGAATTGCACAGTTGGTTGTTCAATTAATGCAACAGTGTATCTCAGTCTCTCTGGGAATG 300

Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 301 GGAAATGACAGTGAAGATGAAGTGAATTCGAGAAGTTGTTATGGAACACATTAGCAAG 119
```

Db 301 GGAATGACAGTGAAGATGAGATTTCGAGAGTGTGTTATCGAACACATTAGCAAG 360
 Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
 Db 361 ATGGAGAAAGATCCATATATATTCAGATCTAGCAAGAAATCTCGTAGATTCAAGCAT 420
 Qy 140 PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
 Db 421 TTTCAAAATTTTCAGTGTGACAACTGATCAAGATTTCGTGATGTTCTTCTCAACTAAGT 480
 Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleGlyAspIleSerLysSerLeu 179
 Db 481 ACCTTGTTCACAGTCCAGGACATGGGAATGCGTAGATGAATACACAGTCCTTA 540
 Qy 180 ValGlyLeuAsnThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
 Db 541 ATAAGTCTGAATACACGCTCTTGAATTTGACCTCTATGTAGAAACACTGAATGTCAA 600
 Qy 200 ValGlnGluAsnAlaPheLysGlnGluGluMetArgLysLeuGluGluArgIleTyr 219
 Db 601 TTCAGAGAGATACACTTAAGGGCAAGAGAAATACAGAAATTAAGAGGCGGTGTCAC 660
 Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
 Db 661 AATCATCAGCAGAAATATGCTCTATGAAAGAAACAGAGTGCAATTTGGAAACAGAAATA 720
 Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
 Db 721 AAAAGAGAAGTGAAGCTCTGAATAACATCACTAATGATCTCAGGCTGAAAGATTGGAA 780
 Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
 Db 781 CATCTCAGAGTGTGAGAAATATCACTTTAATCAAGGT 819

RESULT 13

US-10-733-031-1

; Sequence 1, Application US/10733031

; Publication No. US20040175441A1

; GENERAL INFORMATION:

; APPLICANT: BOOTLAND, LINDA M.

; APPLICANT: BEIFUSS, KATHERINE

; APPLICANT: JILKA, JOSEPH

; APPLICANT: LIZAMA, MALCOLM L.

; APPLICANT: LAMPHEAR, BARRY

; APPLICANT: STREATHFIELD, STEPHEN

; TITLE OF INVENTION: IMMUNIZATION OF FISH WITH PLANT-EXPRESSED RECOMBINANT

; FILE REFERENCE: 10044

; CURRENT APPLICATION NUMBER: US/10733,031

; CURRENT FILING DATE: 2003-12-11

; PRIOR APPLICATION NUMBER: 60/433,381

; PRIOR FILING DATE: 2002-12-13

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn Ver. 3.2

; SEQ ID NO 1

; LENGTH: 459

; TYPE: DNA

; ORGANISM: Hordeum vulgare

US-10-733-031-1

Alignment Scores:

Pred. No.: 4,91e-59 Length: 459

Score: 688.00 Matches: 130

Percent Similarity: 98.5% Conservative: 1

Best Local Similarity: 97.7% Mismatches: 2

Query Match: 33.4% Indels: 0

DB: 7 Gaps: 0

US-10-618-570-2 (1-400) x US-10-733-031-1 (1-459)

Qy 268 ThrLeuLeuGlnGlyAlaArgLysCysSerLeuThrGlyLysThrAsnAspLeuGly 287

Db 58 TCCCTCCGACGCGCGGAGGAGTGTCTCCCTCCGCGAAGTGGACCAATGACCTCGGC 117

Qy 288 SerAsnMetThrIleGlyAlaValAsnSerArgGlyGluPheThrGlyThrTyrIleThr 307
 Db 118 TCCAACATGACCATCGCGCGGTGAATCCAGGGCGGAGTTCCAGGCACCTACATCACC 177
 Qy 308 AlaValThrAlaThrSerAsnGluIleLysGluSerProLeuHisGlyThrGlnAsnThr 327
 Db 178 GCGGTGACCCGACCTCCAAGAGATCAAGAGTCCCTCCACGGTACCCAGAACACC 237
 Qy 328 IleAsnLysArgThrGlnProThrPheGlyPheThrValAsnTrpLysPheSerGluSer 347
 Db 238 ATCAACAAGAGGAGCCAGCCACCTTCGGCTTCACCGTGAACGTGAAGTTCTCCGAGTCC 297
 Qy 348 ThrThrValPheThrGlyGlnCysPheIleAspArgAsnGlyLysGluValLeuLysThr 367
 Db 298 ACCACCGTGTTCACCGCGCAGTGTCTTCATCGACCGCAACGCGCAAGGAGGTCTCAAGACC 357
 Qy 368 MetTrpLeuLeuArgSerSerValAsnAspIleGlyAspLeuTyrLysAlaThrArgVal 387
 Db 358 ATGTGGCTCTCGAGGAGCTCCGTGAATGACATCGCGCAGCATCGAAGGCCACCCGCGTG 417
 Qy 388 GlyIleAsnIlePheThrArgLeuArgThrGlnLysGlu 400
 Db 418 GGCATCAACATCTTCACCCGCTCCGACCCAGGAGGAG 456

RESULT 14

US-11-093-776-6

; Sequence 6, Application US/11093776

; Publication No. US20050172356A1

; GENERAL INFORMATION:

; APPLICANT: Christeller, John Tane

; APPLICANT: Sutherland, Paul William

; APPLICANT: Murray, Colleen

; APPLICANT: Markwick, Ngaire Patricia

; APPLICANT: Phillip, Bruce Allan

; APPLICANT: Malone, Louise Anne

; APPLICANT: Burgess, Elisabeth Phyllis

; APPLICANT: Phung, Margaret Mary

; APPLICANT: Phung, Thai Hong

; APPLICANT: The Horticulture and Food Research Institute of

; TITLE OF INVENTION: Chimeric Polypeptides Allowing Expression of

; FILE REFERENCE: 020829-0001000S

; CURRENT APPLICATION NUMBER: US/11/093,776

; CURRENT FILING DATE: 2005-03-29

; PRIOR APPLICATION NUMBER: US/09/743,690

; PRIOR FILING DATE: 2001-01-12

; PRIOR APPLICATION NUMBER: NZ 331002

; PRIOR FILING DATE: 1998-07-15

; PRIOR APPLICATION NUMBER: WO PCT/NZ99/00110

; PRIOR FILING DATE: 1999-07-15

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 486

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: PPI-I/Avidin

; OTHER INFORMATION: gene fusion

; NAME/KEY: CDS

; LOCATION: (1)..(486)

; OTHER INFORMATION: PPI-I/Avidin fusion protein

US-11-093-776-6

Alignment Scores:

Pred. No.: 5,33e-58 Length: 486

Score: 678.00 Matches: 128

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 32.9% Indels: 0

```
DB: 10 Gaps: 0
US-10-618-570-2 (1-400) x US-11-093-776-6 (1-486)
Qy 273 AlaArgLysCysSerLeuThrGlyLysTrpThrAsnAspLeuGlySerAsnMetThrIle 292
Db 100 GCCAGAAAGTCTCGCTGACTGGGAAATGGACCAACGATCTGGGCTCCACATGACCATC 159
Qy 293 GlyAlaValAsnSerArgGlyGluPheThrGlyThrTyrlleThrAlaValThrAlaThr 312
Db 160 GGGGCTGTGAACAGCAGAGGTGAATTCACAGGCACCTACATCACAGCCGTAAACAGCCACA 219
Qy 313 SerAsnGluIleLysGluSerProLeuHisGlyThrGlnAsnThrIleAsnLysArgThr 332
Db 220 TCNAATGAGATCAAGAGTCAACATTCGATGGGACACAAACACATCAACAGAGACC 279
Qy 333 GlnProThrPheGlyPheThrValAsnTrpLysPheSerGluSerThrThrValPheThr 352
Db 280 CAGCCCACTTTGGCTTCACCGTCAATTTGGAAGTTTTCAGAGTCCACCATGTCTTCACG 339
Qy 353 GlyGlnCysPheIleAspArgAsnGlyLysGluValLeuLysThrMetTrpLeuLeuArg 372
Db 340 GGCAGTGTCTTCATAGACAGAAATGGGAAGAGTCTCTGAAGACCATGTGGCTCTCGG 399
Qy 373 SerSerValAsnAspIleGlyAspAspTrpLysAlaThrArgValGlyIleAsnIlePhe 392
Db 400 TCNAGTGTATGACATTTGGTGTGATGACTGGAAAGCTACCAAGGTCGGCATCATCTTTC 459
Qy 393 ThrArgLeuArgThrGlnLysGlu 400
Db 460 ACTCGCTCGGCACACAGAGGAG 483
```

RESULT 15

```
US-11-093-776-2
; Sequence 2, Application US/11093776
; Publication No. US20050172356A1
; GENERAL INFORMATION:
; APPLICANT: Christeller, John Tane
; APPLICANT: Sutherland, Paul William
; APPLICANT: Murray, Colleen
; APPLICANT: Markwick, Ngaire Patricia
; APPLICANT: Philip, Bruce Allan
; APPLICANT: Malone, Louise Anne
; APPLICANT: Burgess, Elisabeth Phyllis
; APPLICANT: Phung, Margaret Mary
; APPLICANT: The Horticulture and Food Research Institute of
; APPLICANT: New Zealand Limited
; TITLE OF INVENTION: Chimeric Polypeptides Allowing Expression of
; FILE REFERENCE: 020829-000100US
; CURRENT APPLICATION NUMBER: US/11/093,776
; PRIOR FILING DATE: 2005-03-29
; PRIOR FILING DATE: 2001-01-12
; PRIOR FILING DATE: 1998-07-15
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:avdin cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (44)..(502)
; OTHER INFORMATION: avdin (pGEMav)
; FEATURE:
```

```
; NAME/KEY: sig peptide
; LOCATION: (44)..(115)
; OTHER INFORMATION: signal sequence
US-11-093-776-2
```

```
Alignment Scores:
Pred. No.: 6.8e-58 Length: 576
Score: 678.00 Matches: 128
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 32.9% Indels: 0
DB: 10 Gaps: 0
```

US-10-618-570-2 (1-400) x US-11-093-776-2 (1-576)

```
Qy 273 AlaArgLysCysSerLeuThrGlyLysTrpThrAsnAspLeuGlySerAsnMetThrIle 292
Db 116 GCCAGAAAGTCTCGCTGACTGGGAAATGGACCAACGATCTGGGCTCCACATGACCATC 175
Qy 293 GlyAlaValAsnSerArgGlyGluPheThrGlyThrTyrlleThrAlaValThrAlaThr 312
Db 176 GGGGCTGTGAACAGCAGAGGTGAATTCACAGGCACCTACATCACAGCCGTAAACAGCCACA 235
Qy 313 SerAsnGluIleLysGluSerProLeuHisGlyThrGlnAsnThrIleAsnLysArgThr 332
Db 236 TCNAATGAGATCAAGAGTCAACATTCGATGGGACACAAACACATCAACAGAGACC 295
Qy 333 GlnProThrPheGlyPheThrValAsnTrpLysPheSerGluSerThrThrValPheThr 352
Db 296 CAGCCCACTTTGGCTTCACCGTCAATTTGGAAGTTTTCAGAGTCCACCATGTCTTCACG 355
Qy 353 GlyGlnCysPheIleAspArgAsnGlyLysGluValLeuLysThrMetTrpLeuLeuArg 372
Db 356 GGCAGTGTCTTCATAGACAGAAATGGGAAGAGTCTCTGAAGACCATGTGGCTCTCGG 415
Qy 373 SerSerValAsnAspIleGlyAspAspTrpLysAlaThrArgValGlyIleAsnIlePhe 392
Db 416 TCNAGTGTATGACATTTGGTGTGATGACTGGAAAGCTACCAAGGTCGGCATCATCTTTC 475
Qy 393 ThrArgLeuArgThrGlnLysGlu 400
Db 476 ACTCGCTCGGCACACAGAGGAG 499
```

Search completed: February 27, 2006, 10:42:01
Job time : 1037 secs

GenCore version 5.1.7

Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 27, 2006, 09:29:22 ; Search time 507 Seconds

(without alignments)

1682.341 Million cell updates/sec

Title: US-10-618-570-2

Perfect score: 2059

Sequence: 1 MAQWDDPPDQEDTDSCTES.....DWKATRVGINFTRLRTQKE 400

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 7209121 seqs, 1066183437 residues

Total number of hits satisfying chosen parameters: 14418242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abss/ABSSWEB spool/US10618570/runat 27022006 063904 13261/app query.fasta_1
-DB=Published Applications NA New -QFMT=fastap -SUFFIX=rnpbn -MINMATCH=0.1
-LOOFCLO=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs05p
-USER=US10618570 @CGN 1 1 335 @runat 27022006 063904 13261 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

Database : Published Applications NA New:

1:	/cgn2_6/ptodata/1/pubpna/US08 NEW PUB.seq.*
2:	/cgn2_6/ptodata/1/pubpna/US06 NEW PUB.seq.*
3:	/cgn2_6/ptodata/1/pubpna/US07 NEW PUB.seq.*
4:	/cgn2_6/ptodata/1/pubpna/PCT NEW PUB.seq.*
5:	/cgn2_6/ptodata/1/pubpna/US09 NEW PUB.seq.*
6:	/cgn2_6/ptodata/1/pubpna/US09 NEW PUB.seq.*
7:	/cgn2_6/ptodata/1/pubpna/US10 NEW PUB.seq.*
8:	/cgn2_6/ptodata/1/pubpna/US10 NEW PUB.seq.*
9:	/cgn2_6/ptodata/1/pubpna/US11 NEW PUB.seq.*
10:	/cgn2_6/ptodata/1/pubpna/US11 NEW PUB.seq.*
11:	/cgn2_6/ptodata/1/pubpna/US11 NEW PUB.seq.*
12:	/cgn2_6/ptodata/1/pubpna/US11 NEW PUB.seq.*
13:	/cgn2_6/ptodata/1/pubpna/US60 NEW PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1016.5	49.4	2877	8	US-10-995-561-372 Sequence 372, App
2	1016.5	49.4	3336	8	US-10-995-561-370 Sequence 370, App
3	1016.5	49.4	3525	8	US-10-995-561-369 Sequence 369, App
4	1016.5	49.4	3719	8	US-10-995-561-371 Sequence 371, App

5	1016.5	49.4	3791	8	US-10-995-561-368 Sequence 368, App
6	480.5	23.8	119036	8	US-10-995-561-13314 Sequence 13314, A
7	282	13.7	201	8	US-10-995-561-9974 Sequence 9974, Ap
8	282	13.7	201	8	US-10-995-561-10011 Sequence 10011, A
9	282	13.7	201	8	US-10-995-561-10046 Sequence 10046, A
10	282	13.7	201	8	US-10-995-561-10086 Sequence 10086, A
11	282	13.7	201	8	US-10-995-561-10109 Sequence 10109, A
12	273	13.3	201	8	US-10-995-561-9940 Sequence 9940, Ap
13	273	13.3	201	8	US-10-995-561-9977 Sequence 9977, Ap
14	273	13.3	201	8	US-10-995-561-10014 Sequence 10014, A
15	273	13.3	201	8	US-10-995-561-10049 Sequence 10049, A
16	273	13.3	201	8	US-10-995-561-10100 Sequence 10100, A
17	272	13.2	201	8	US-10-995-561-9941 Sequence 9941, Ap
18	272	13.2	201	8	US-10-995-561-9978 Sequence 9978, Ap
19	272	13.2	201	8	US-10-995-561-10015 Sequence 10015, A
20	272	13.2	201	8	US-10-995-561-10051 Sequence 10051, A
21	272	13.2	201	8	US-10-995-561-10101 Sequence 10101, A
22	263	12.8	201	8	US-10-995-561-9939 Sequence 9939, Ap
23	263	12.8	201	8	US-10-995-561-9976 Sequence 9976, Ap
24	263	12.8	201	8	US-10-995-561-10013 Sequence 10013, A
25	263	12.8	201	8	US-10-995-561-10048 Sequence 10048, A
26	263	12.8	201	8	US-10-995-561-10097 Sequence 10097, A
27	254	12.3	201	8	US-10-995-561-9973 Sequence 9973, Ap
28	254	12.3	201	8	US-10-995-561-10010 Sequence 10010, A
29	254	12.3	201	8	US-10-995-561-10045 Sequence 10045, A
30	254	12.3	201	8	US-10-995-561-10085 Sequence 10085, A
31	254	12.3	201	8	US-10-995-561-10108 Sequence 10108, A
32	241	11.7	201	8	US-10-995-561-9956 Sequence 9956, Ap
33	241	11.7	201	8	US-10-995-561-9993 Sequence 9993, Ap
34	241	11.7	201	8	US-10-995-561-10029 Sequence 10029, A
35	241	11.7	201	8	US-10-995-561-10068 Sequence 10068, A
36	241	11.7	201	8	US-10-995-561-10092 Sequence 10092, A
37	241	11.7	201	8	US-10-995-561-52999 Sequence 52999, A
38	233	11.3	286	12	US-11-128-061-1104 Sequence 1104, Ap
39	233	11.3	286	12	US-11-128-061-4746 Sequence 4746, Ap
40	233	11.3	286	12	US-11-128-049-1104 Sequence 1104, Ap
41	233	11.3	286	12	US-11-128-049-4746 Sequence 4746, Ap
42	220.5	10.7	201	8	US-10-995-561-9947 Sequence 9947, Ap
43	220.5	10.7	201	8	US-10-995-561-9985 Sequence 9985, Ap
44	220.5	10.7	201	8	US-10-995-561-10020 Sequence 10020, A
45	220.5	10.7	201	8	US-10-995-561-10059 Sequence 10059, A

ALIGNMENTS

RESULT 1

US-10-995-561-372

; Sequence 372, Application US/10995561

; Publication No. US20050272054A1

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF

; TITLE OF INVENTION: DETECTION AND USES THEREOF

; FILE REFERENCE: CLO01559

; CURRENT APPLICATION NUMBER: US/10/995.561

; CURRENT FILING DATE: 2004-11-24

; NUMBER OF SEQ ID NOS: 85702

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 372

; LENGTH: 2877

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-995-561-372

Alignment Scores:

Pred. No.:	2.32e-101	Length:	2877
Score:	1016.50	Matches:	193
Percent Similarity:	85.3%	Conservative:	40
Best Local Similarity:	70.7%	Mismatches:	39
Query Match:	49.4%	Indels:	1
DB:	8	Gaps:	1

```
US-10-618-570-2 (1-400) x US-10-995-561-372 (1-2877)
QY 1 MetAlaGlnTrpAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
Db 175 ATGGAGCAGTGGGATCACTTTACAAATCAACAGGAGGACACTGATAGCTGCTCCGAATCT 234
QY 21 VallysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 235 GTGAATAATTGATCTCGCTCAATGACAGCTTTGCTTCTCCGAATCTAAACACAGCCCT 294
QY 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuLeuLeuVal 60
Db 295 TMCCTTCAAGAGAACTCAAGTCTTCAAGCTGCACTGRTTGGCCCTTACCTCTCGTG 354
QY 61 PheValValLeuValProIleGlyLeuValAlaAlaGlnLeuLeuLysTrpGluThr 80
Db 355 TTTGAGTTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCCTGAAGTGGGAACG 414
QY 81 LysAsnCysThrValGlySerValAsnAla---AspLysSerProSerProGluGlyLys 99
Db 415 AAGAAATTCCTCAGTTAGTTCAGTCAATGCAATGATATACTCAAGTCTCACGGGAAA 474
QY 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 475 GGAATGACAGCAGAAAGAAATGAGATTTCAAGAAAGCTTTATGGAACACATGAGCAAC 534
QY 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 535 ATGGAGAAGAGAAATCCAGCATATTTAGACATGGAAGCCCACTCATGGACACAGAGCAT 594
QY 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 595 TTCCAAATTTCCAGCATGACCACTGATCAAGATTTAATGACATCTCTCGACGTAAGT 654
QY 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspLysSerLysSerLeu 179
Db 655 ACCTTGTTTCTCAGTCAGGACATCGGAATGCAATKATKATGAAATCTCCAAGTCTCTTA 714
QY 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 715 ATAAGTTGAATACCAATGCTGATTTGAGCTCAACATAGAAAATCTCAATGGCAAA 774
QY 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 775 ATCCAGAGAAATACCTTCAACACACAGGAAATCAGTAATTAGAGGACGCTGTTTAC 834
QY 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGlnGluIle 239
Db 835 AATGTATCRGAGAAATATGCTATGAAAGAAACAAAGTGCATTTGGAACAGGAAATA 894
QY 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
Db 895 AAAGGAGAAGTCAAAAGTACTGAATACATCACTAATGATCTCAGACTGAAAGATTGGGA 954
QY 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 955 CATTCCTCAGACCTTGAGAAATATCAATTTAATCAAGT 993

RESULT 2
US-10-995-561-370
; Sequence 370, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 370
```

```
; Sequence 369, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995.561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 369
; LENGTH: 3525
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-369

Alignment Scores:
Pred. No.: 3,03e-101 Length: 3525
Score: 1016.50 Matches: 193
Percent Similarity: 85.3% Conservative: 40
Best Local Similarity: 70.7% Mismatches: 39
Query Match: 49.4% Indels: 1
DB: 8 Gaps: 1

US-10-618-570-2 (1-400) x US-10-995-561-369 (1-3525)

Qy 1 MetAlaGlnTTPAspAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
Db 160 ATGAGCAGTGGGATCACTTTTCACATCAACAGGAGGACACTGATAGCTGCTCGAATCT 219
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 220 GTGAAATTTGATGCTGCTCAATGACAGCTTTGCTTCTCGAATCTCAAAACACAGCCT 279
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuLeuLeuVal 60
Db 280 TMCCTTCAAGAGAARCTCAAGCTCTCAAGCTGCACTGRTTGGCCTTTACCTCTCGTG 339
Qy 61 PheValValLeuValProIleGlyLeuValAlaAlaGlnLeuLeuLysThrGluThr 80
Db 340 TTTCAGATTCTCACTCCCTCTCATTTGGAATAGTGGAGCTCAACTCTCGAATGGGAACG 399
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspLysSerProSerProGluGlyLys 99
Db 400 AAGATTGCTCAGTTAGTTCAACTAATGCAATATGATATACTCAAGTCTCAACGGAAA 459
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 460 GGAATGACACGACGAAAGAAATGAGATTTCAAGAAAGCTTTTATGGAACACATGAGCAAC 519
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 520 ATGGAGAAGAAATCCAGCATATTTAGCATGGAAGCCCACTCATGGAACACAGAGCAT 579
Qy 140 PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 580 TTCCAAATTTTCAGATGACACTGATCAAGATTTAATGATATCTTCTGACGTAAGT 639
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleGlyAspLysSerLysSerLeu 179
Db 640 ACCTTGTTTCTCAGTCAGGACATGGGAATGCAATKATAGAAATCTCAAGTCCCTTA 699
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 700 ATAAAGTTGAATACCATCTGCTGATTTGAGCTCAACATAGAAATCTGAATGGCAAA 759
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 760 ATCCAGAGATATACCTTCAACACACAGAGAAATCAGTAATTAGAGGCGGTGTTC 819
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGlnGlnGluIle 239
; Sequence 371, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995.561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 371
; LENGTH: 3719
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-371

Alignment Scores:
Pred. No.: 3,25e-101 Length: 3719
Score: 1016.50 Matches: 193
Percent Similarity: 85.3% Conservative: 40
Best Local Similarity: 70.7% Mismatches: 39
Query Match: 49.4% Indels: 1
DB: 8 Gaps: 1

US-10-618-570-2 (1-400) x US-10-995-561-371 (1-3719)

Qy 1 MetAlaGlnTTPAspAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
Db 175 ATGGAGCAGTGGGATCACTTTTCACATCAACAGGAGGACACTGATAGCTGCTCGAATCT 234
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 235 GTGAAATTTGATGCTGCTCAATGACAGCTTTGCTTCTCGAATCTCAAAACACAGCCT 294
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuLeuLeuVal 60
Db 295 TMCCTTCAAGAGAARCTCAAGCTCTCAAGCTGCACTGRTTGGCCTTTACCTCTCGTG 354
Qy 61 PheValValLeuValProIleGlyLeuValAlaAlaGlnLeuLeuLysThrGluThr 80
Db 355 TTTCAGATTCTCACTCCCTCTCATTTGGAATAGTGGAGCTCAACTCTCGAATGGGAACG 414
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspLysSerProSerProGluGlyLys 99
Db 415 AAGATTGCTCAGTTAGTTCAACTAATGCAATATGATATACTCAAGTCTCAACGGAAA 474
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 475 GGAATGACACGACGAAAGAAATGAGATTTCAAGAAAGCTTTTATGGAACACATGAGCAAC 534
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 535 ATGGAGAAGAAATCCAGCATATTTTAGACATGGAAGCCCACTCATGGAACACAGAGCAT 594
Qy 140 PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 595 TTCCAAATTTTCAGCATGACACTGATCAAGATTTAATGACATCTTCTGACGTAAGT 654
```


Db 52113 CGGAATGATCATATGGGATCGTGTATGCTTCAAGGATCAGCCCATGACTGGT 52172
Qy 56 LeuTyrLeuIleValPheValValLeuValProIleIleGlyLeValAlaAlaGlnLeu 75
Db 52173 TTAATTTTACTTAATAT-----ACAGCTCAACTC 52202
Qy 76 LeuLysTrpGluThrLysAsnCySerValGlySerValAsnAla---AspLeuSerPro 94
Db 52203 CTGAAGTGGGAACCAAGAAATGCTCAGTTAGTCACTAATGCAATATGATATAACTCAA 52262
Qy 95 SerProGluGlyLysGlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMet 114
Db 52263 AGTCTCACGGGAAGGAATACACAGCRAAGGAATGAGATTTCAAGAAGCTTTATG 52322
Qy 115 GluArgMetSerAsnMetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeu 134
Db 52323 GAACACATGAGCAACATGAGAGAGAAATCCAGCATATTTTAGACATGGAAGCCAACTC 52382
Qy 135 LeuAspAlaLysAsnPheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspVal 154
Db 52383 ATGGACACAGAGCATTTCCAAATTTTCAGCATGACACTGATCAAGATTTTAATGACATT 52442
Qy 155 LeuPheGlnLeuAsnSerLeuSerSerIleGlnGluHisGluAsnIleIleGlyAsp 174
Db 52443 CTCTCGCAGCTAAGTACCTTGTTCCTCAGTCAGGACATGGGAATGCATATKATGAA 52502
Qy 175 IleSerLysSerLeuValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGlu 194
Db 52503 ATCTCCAAGTCCTTAATAGTTGAATACCACTATGCTTGTATTTGCAGCTCAACATAGAA 52562
Qy 195 ThrLeuAsnGlyArgValGluAsnAlaPheLysGlnGluAsnMetArgLysLeu 214
Db 52563 AATCTGAATGCAAAATCCAGAGAAATACCTTCAACAAACAA----- 52604
Qy 215 GluGluArgIleTyrAsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyr 234
Db 52605 GAGGTAGAGTTTGGAGTCAAGTCACTGTGGAAGTCAATCCCAAGTCTCGT 52664
Qy 235 LeuGluGlnGluIleLysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAsp----- 252
Db 52665 CTCCTTAGGCAGTACAGG---AAAGCAAAAGTTGCAAACTGTTTCATCATCTGTTTA 52721
Qy 253 -----LeuArgLeuLysAspTrpGluHisSer---GlnThrLeuLysAsn 266
Db 52722 CCCTGATTTCTATGCTCACTATCCAGATTTTAAACGAATGTTAAGGAGTTTAAATGG 52781
Qy 267 IleThrLeuLeuGlnGlyAlaArgLysCySerLeuThrGlyLysTrpThrAsnAspLeu 286
Db 52782 ATTCGCTTCTCTAAACTCTTAGAAA----- 52808
Qy 287 GlySerAsnMetThrIleGlyAlaValAsnSerArgGlyGluPheThrGlyThrTyrIle 306
Db 52809 -----AGGGGAATTTTCCCATGGAATATAG 52835
Qy 307 ThrAlaValThrAlaThrSerAsnGluIleLysGluSer 319
Db 52836 ACTCCCTTACATATTTTATAAATACTGTTAAATGTCT 52874

RESULT 7

US-10-995-561-9974
; Sequence 9974, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; NUMBER OF SEQ ID NOS: 2004-11-24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9974

LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-9974
Alignment Scores:
Pred. No.: 7.5e-22 Length: 201
Score: 282.00 Matches: 52
Percent Similarity: 93.9% Conservativeness: 10
Best Local Similarity: 78.8% Mismatches: 4
Query Match: 13.7% Indels: 0
DB: 8 Gaps: 0
US-10-618-570-2 (1-400) x US-10-995-561-9974 (1-201)
Qy 13 AspThrAspSerCySerThrGluSerValLysPheAspAlaArgSerValThrAlaLeuLeu 32
Db 3 GACACTGATAGTCTCGGAATCTGTGAAATTTGATGCTCGCTCAATGACAGCTTTGCTT 62
Qy 33 ProProHisProLysAsnGlyProThrLeuGlnGluArgMetLysSerTyrLysThrAla 52
Db 63 CCTCCGAATCCTAAAAACAGCCCTTCCCTTCAAGAGAACTGAAAGTCTTCAAGAGCTGCA 122
Qy 53 LeuIleThrLeuTyrLeuIleValPheValValLeuValProIleIleGlyLeValAla 72
Db 123 CTGATGCCCCTTTACCTCCTCGTGTTCGAGTTTCTCATCCCTCTCATTTGGAAATAGTGCA 182
Qy 73 AlaGlnLeuLeuLysTrp 78
Db 183 GCTCAACTCCTCGAAGTGG 200
RESULT 8
US-10-995-561-10011
; Sequence 10011, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10011
LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-10011
Alignment Scores:
Pred. No.: 7.5e-22 Length: 201
Score: 282.00 Matches: 52
Percent Similarity: 93.9% Conservativeness: 10
Best Local Similarity: 78.8% Mismatches: 4
Query Match: 13.7% Indels: 0
DB: 8 Gaps: 0
US-10-618-570-2 (1-400) x US-10-995-561-10011 (1-201)
Qy 13 AspThrAspSerCySerThrGluSerValLysPheAspAlaArgSerValThrAlaLeuLeu 32
Db 3 GACACTGATAGTCTCGGAATCTGTGAAATTTGATGCTCGCTCAATGACAGCTTTGCTT 62
Qy 33 ProProHisProLysAsnGlyProThrLeuGlnGluArgMetLysSerTyrLysThrAla 52
Db 63 CCTCCGAATCCTAAAAACAGCCCTTCCCTTCAAGAGAACTGAAAGTCTTCAAGAGCTGCA 122
Qy 53 LeuIleThrLeuTyrLeuIleValPheValValLeuValProIleIleGlyLeValAla 72
Db 123 CTGATGCCCCTTTACCTCCTCGTGTTCGAGTTTCTCATCCCTCTCATTTGGAAATAGTGCA 182

```
Qy 73 AlaGlnLeuLeuLysTyr 78
|||
Db 183 GCTCAACTCTCTGAAGTGG 200

RESULT 9
US-10-995-561-10046
; Sequence 10046, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10046
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-10046

Alignment Scores:
Pred. No.: 7.5e-22 Length: 201
Score: 282.00 Matches: 52
Percent Similarity: 93.9% Conservative: 10
Best Local Similarity: 78.8% Mismatches: 4
Query Match: 13.7% Indels: 0
DB: 8 Gaps: 0

US-10-618-570-2 (1-400) x US-10-995-561-10046 (1-201)
Qy 13 AspThrAspSerCysThrGluSerValLysPheAspAlaArgSerValThrAlaLeuLeu 32
|||
Db 3 GACACTGATAGTCTCCGAATCTGTGAATTTGATGCTCGCTCAATGACAGCTTTGCTT 62

Qy 33 ProProHisProLysAsnGlyProThrLeuGlnGluArgMetLysSerTyrLysThrAla 52
|||
Db 63 CTTCCGAATCTTAAACAGCCCTTCCCTTCAAGAGAARCTGAAGTCTCTCAAGAGCTGCA 122

Qy 53 LeuIleThrLeuTyrLeuIleValPheValValLeuValProIleIleGlyIleValAla 72
|||
Db 123 CTGATTGCCCTTTTACCTCTCGTGTTCGAGTTTCTCATCCCTCTCTCATTTGGAATAGTGCA 182

Qy 73 AlaGlnLeuLeuLysTyr 78
|||
Db 183 GCTCAACTCTCTGAAGTGG 200

RESULT 10
US-10-995-561-10086
; Sequence 10086, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10086
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-10086

Alignment Scores:
Pred. No.: 7.5e-22 Length: 201
Score: 282.00 Matches: 52
Percent Similarity: 93.9% Conservative: 10
Best Local Similarity: 78.8% Mismatches: 4
Query Match: 13.7% Indels: 0
DB: 8 Gaps: 0

US-10-618-570-2 (1-400) x US-10-995-561-10086 (1-201)
Qy 13 AspThrAspSerCysThrGluSerValLysPheAspAlaArgSerValThrAlaLeuLeu 32
|||
Db 3 GACACTGATAGTCTCCGAATCTGTGAATTTGATGCTCGCTCAATGACAGCTTTGCTT 62

Qy 33 ProProHisProLysAsnGlyProThrLeuGlnGluArgMetLysSerTyrLysThrAla 52
|||
Db 63 CTTCCGAATCTTAAACAGCCCTTCCCTTCAAGAGAARCTGAAGTCTCTCAAGAGCTGCA 122

Qy 53 LeuIleThrLeuTyrLeuIleValPheValValLeuValProIleIleGlyIleValAla 72
|||
Db 123 CTGATTGCCCTTTTACCTCTCGTGTTCGAGTTTCTCATCCCTCTCTCATTTGGAATAGTGCA 182

Qy 73 AlaGlnLeuLeuLysTyr 78
|||
Db 183 GCTCAACTCTCTGAAGTGG 200

RESULT 11
US-10-995-561-10109
; Sequence 10109, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10109
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-10109

Alignment Scores:
Pred. No.: 7.5e-22 Length: 201
Score: 282.00 Matches: 52
Percent Similarity: 93.9% Conservative: 10
Best Local Similarity: 78.8% Mismatches: 4
Query Match: 13.7% Indels: 0
DB: 8 Gaps: 0

US-10-618-570-2 (1-400) x US-10-995-561-10109 (1-201)
Qy 13 AspThrAspSerCysThrGluSerValLysPheAspAlaArgSerValThrAlaLeuLeu 32
|||
Db 3 GACACTGATAGTCTCCGAATCTGTGAATTTGATGCTCGCTCAATGACAGCTTTGCTT 62

Qy 33 ProProHisProLysAsnGlyProThrLeuGlnGluArgMetLysSerTyrLysThrAla 52
|||
Db 63 CTTCCGAATCTTAAACAGCCCTTCCCTTCAAGAGAARCTGAAGTCTCTCAAGAGCTGCA 122

Qy 53 LeuIleThrLeuTyrLeuIleValPheValValLeuValProIleIleGlyIleValAla 72
|||
Db 123 CTGATTGCCCTTTTACCTCTCGTGTTCGAGTTTCTCATCCCTCTCTCATTTGGAATAGTGCA 182

Qy 73 AlaGlnLeuLeuLysTyr 78
|||
Db 183 GCTCAACTCTCTGAAGTGG 200

RESULT 12
US-10-995-561-9940
; Sequence 9940, Application US/10995561
; Publication No. US20050272054A1
```



```
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10049
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-10049
```

```
Alignment Scores:
Pred. No.: 7.38e-21 Length: 201
Score: 273.00 Matches: 51
Percent Similarity: 92.4% Conservative: 10
Best Local Similarity: 77.3% Mismatches: 5
Query Match: 13.3% Indels: 0
DB: 8 Gaps: 0
```

```
US-10-618-570-2 (1-400) x US-10-995-561-10049 (1-201)
```

```
Qy 9 AspGlnGlnGluAspThrAspSerCysThrGluSerValLysPheAspAlaArgSerVal 28
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 4 AATCAACAGGAGGACACTGATAGCTGCTCCGAATCTGGAATTTGATGCTCGCTCAATG 63

Qy 29 ThrAlaLeuLeuProHisProLysAsnGlyProThrLeuGlnGluArgMetLysSer 48
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 64 ACAGCTTTGCTTCCTCCGAATCTTAAAAACAGCCCTTWCCTTCAAGAGAACTGAAGTCC 123

Qy 49 TyrLysThrAlaLeuIleThrLeuTyrLeuIleValPheValValLeuValProIleIle 68
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 124 TTCAAAGCTGCACCTGATTTGCCCTTTACCTCTCGTGTTCAGTTTCATCCCTCTCAT 183

Qy 69 GlyIleValAlaAlaGln 74
   ::::::::::::::::::::
Db 184 GGATAGTGGCAGCTCAA 201
```

```
Search completed: February 27, 2006, 09:49:08
Job time : 545 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 27, 2006, 09:22:16 ; Search time 230 Seconds
(without alignments)
3091.412 Million cell updates/sec

Title: US-10-618-570-2

Perfect score: 2059

Sequence: 1 MAQWDDPPDQEDTDSCTES.....DWKATRVGINFTLRITQKE 400

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abss/ABSSWEB gpool/US10618570/runat 27022006 063859 13138/app query.fasta_1
-DB=Issued Patents NA -QWMT=fastap -SUFFIX=rni -MINWATCH=0.1 -LOOPECL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss03p -USER=US10618570 @CGSN 1 1 290 @runat 27022006 063859 13138
-NCPU=6 -ICPU=3 -NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1381.5	67.1	1588	9 5510466-3
2	1381.5	67.1	1807	9 Patent No. 5510466
3	1042.5	50.6	1366	3 US-09-949-016-3387
4	1042.5	50.6	1367	2 US-08-453-117-3
5	1042.5	50.6	1367	2 US-08-948-222-3
6	1042.5	50.6	1367	2 US-08-973-145-3
7	1042.5	50.6	1367	3 US-09-016-434-1260
8	1042.5	50.6	1367	6 PCT-US96-08081-3
9	1042.5	50.6	2028	2 US-08-453-117-1

10	1042.5	50.6	2028	2	US-08-948-222-1	Sequence 1, Appli
11	1042.5	50.6	2028	2	US-08-973-145-1	Sequence 1, Appli
12	1042.5	50.6	2028	6	PCT-US96-08081-1	Sequence 1, Appli
13	1039.5	50.5	2037	2	US-08-154-365-1	Sequence 1, Appli
14	688	33.4	484	2	US-08-554-586-1	Sequence 1, Appli
15	678	32.9	604	3	US-08-831-399-3	Sequence 3, Appli
16	678	32.9	604	3	US-08-366-882-3	Sequence 3, Appli
17	678	32.9	604	3	US-09-368-772-3	Sequence 3, Appli
18	495.5	24.1	55886	3	US-09-949-016-15129	Sequence 15129, A
19	275	13.4	601	3	US-09-949-016-121418	Sequence 121418, A
20	173	8.4	245	3	US-09-513-999C-36319	Sequence 36319, A
21	156.5	7.6	601	3	US-09-949-016-121413	Sequence 121413, A
22	147.5	7.2	1173	3	US-09-142-974B-4	Sequence 4, Appli
23	145.5	7.1	1176	3	US-09-142-974B-3	Sequence 3, Appli
24	138.5	6.7	498	3	US-09-117-447-7	Sequence 7, Appli
25	138.5	6.7	525	2	US-07-924-028A-2	Sequence 2, Appli
26	138.5	6.7	552	6	PCT-US93-05240-13	Sequence 13, Appli
27	138.5	6.7	638	3	US-08-831-399-1	Sequence 1, Appli
28	138.5	6.7	638	3	US-08-381-430-1	Sequence 1, Appli
29	138.5	6.7	638	3	US-09-366-862-1	Sequence 1, Appli
30	138.5	6.7	638	3	US-09-368-772-1	Sequence 1, Appli
31	138.5	6.7	1131	9	5168049-1	Patent No. 5168049
32	138.5	6.7	1266	3	US-09-938-270B-2	Sequence 2, Appli
33	138.5	6.7	1296	2	US-08-491-988-6	Sequence 6, Appli
34	138.5	6.7	1356	2	US-08-491-988-4	Sequence 4, Appli
35	135.5	6.6	384	3	US-08-831-399-15	Sequence 15, Appli
36	135.5	6.6	384	3	US-09-366-862-15	Sequence 15, Appli
37	135.5	6.6	384	3	US-09-368-772-15	Sequence 15, Appli
38	135.5	6.6	387	2	US-08-211-833-1	Sequence 1, Appli
39	135.5	6.6	387	2	US-08-434-718-1	Sequence 1, Appli
40	135.5	6.6	1257	2	US-08-491-988-8	Sequence 8, Appli
41	135.5	6.6	30549	3	US-09-134-001C-322	Sequence 322, App
42	132.5	6.4	354	3	US-07-780-717C-6	Sequence 6, Appli
43	129.5	6.3	4407	3	US-09-949-016-1690	Sequence 1690, Ap
44	128.5	6.2	7425	3	US-09-453-702B-212	Sequence 212, App
45	128.5	6.2	7425	3	US-10-114-170-212	Sequence 212, App

ALIGNMENTS

RESULT 1
5510466-3
; Patent No. 5510466
; APPLICANT: KREIGER, MONTY;KODAMA, TATSUHIKO
; TITLE OF INVENTION: SCAVENGER RECEPTOR PROTEIN AND ANTIBODY
; THERETO
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,400
; FILING DATE: 16-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 997,113
; FILING DATE: 24-DEC-1992
; APPLICATION NUMBER: 391,486
; FILING DATE: 09-AUG-1989
; APPLICATION NUMBER: 272,002
; FILING DATE: 15-NOV-1988
; SEQ ID NO:3:
; LENGTH: 1588
5510466-3

Alignment Scores:
Pred. No.: 4.63e-154 Length: 1588
Score: 1381.50 Matches: 307
Percent Similarity: 79.4% Conservatives: 24
Best Local Similarity: 73.8% Mismatches: 66
Query Match: 7.1% Indels: 22
DB: 9 Gaps: 6

US-10-618-570-2 (1-400) x 5510466-3 (1-1588)

Qy 1 MetalaGlnTrpAspAppPheProAspGlnGluAspThrAppSerCysThrGluSer 20
|||||

Db 4 ATGGCACAGTGGGATGACTTTCTCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCT 63
Qy 21 VallysPheAspAlaAatqSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 64 GTGAAGTTCGATGCTCGCTCAGTACAGACTTTGCTCTCCCTCAATCTTAAATAATGGCCCA 123
Qy 41 ThrLeuGlnGluArgMetLysSerTyrlsThrAlaLeuLeuThrLeuTyrlsVal 60
Db 124 ACTCTCAAGAGAGAGTGAAGTCTTATAAAGTGCACCTGATCACCCTTTATCTCATTTG 183
Qy 61 PheValValLeuValProIlelleGlyIleValAlaAlaGlnLeuLeuTyrlsGluThr 80
Db 184 TTTGTAGTTCCTGTCCTCCCATCTTGGCATAGTGGCAGCTCAGCTCTCCCTCAATGGGAAACG 243
Qy 81 LysAsnCythrValGlySerValAsnAlaAspIleSerProSerProGluGlyGly 100
Db 244 AAGAATTGCACGGTTGGCTCAGTTAATGCAGATATATCTCAAGTCCGGAAGGCAAGGA 303
Qy 101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120
Db 304 AATGGCAGTGAAGATGAATGAGATTTCCAGAGCTGTGATGGAAAGCATGAGCAACATG 363
Qy 121 GluSerArgIleGlnTyLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPhe 140
Db 364 GAAAGCAATCCAGTATCTTTCCAGATTAATGAAGCCAAATCTCTAGATGCTAAGAAATTC 423
Qy 141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
Db 424 CAAAATTTCCAGCATTAACCACTGATCAAGATTTAATGATGTTCTTTTCCAGCTAAATTC 483
Qy 161 LeuLeuSerSerIleGlnGluHisGluAsnIlelleGlyAspIleSerLysSerLeuVal 180
Db 484 TTACTTTCTCTCCATCCAGAAACATGAGAAATATCATAGGGATATCTCCAGCTATTAGTA 543
Qy 181 GlyLeuAsnThrThrValLeuAspLeuGlnPheSerIlelleGluThrLeuAsnGlyArgVal 200
Db 544 GGTCTGAACACCAACAGTACTTGTATGTCAGTTCAGTATTGAACACACTGAATGGCAGAGTC 603
Qy 201 GlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyAsn 220
Db 604 CAAGAGAAATGCAATTTAAACCAACAGAGAGATGCGTAAATTAGAGGAGCGTATATACAA 663
Qy 221 AlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyLeuGluGlnGluIleLys 240
Db 664 GCATCAGCAGAAATTAAGTCTCTAGATCAAAACCAAGTATATTTGGAACAGAAATAAAA 723
Qy 241 GlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGluHis 260
Db 724 GGGGAAATGAACCTGTTGAATAATATCACTAATGATCTGAGGCTCAAGGATTGGGAACAT 783
Qy 261 SerGlnThrLeuLysAsnIleThrLeuLeuGlnGlyAla----- 273
Db 784 TCTCAGACATTTGAAAAATATCACTTTACTCAAGGTCCTCTCGGACCTCCAGGTGCAAAAA 843
Qy 274 -----ArgLysCythrLeuGlyLysTrpThrAsnAspLeuGly---SerAsnMet 290
Db 844 GGAGATAGAGGCCCTCTCTGGACA---AATGGTATACAGGCTTTCCAGCTCTAATAGG 899
Qy 291 ThrIleGlyAlaValAsnSerArgGlyGluPheThrGlyThrTyrlsIleThrAlaValThr 310
Db 900 TACTCCAGTCTTAAAGGTGATCGGGGATCTCTGTTTACTCTGGAGTTCGAGGATTCCTCC 959
Qy 311 AlaThrSerAsnGluIleLysGluSer-ProLeuHisGlyThrGlnAsnThrIleAsnLys 330
Db 960 AGGACCAATGGGAAGACCGGAGCCAGGACTTAATGGCAAAAAGCCAGAGAGGAGAGA 1019
Qy 330 eArgThrGlnProThrPhe-----GlyPheThrVal-AsnTrpLysPheSerGlu 347
Db 1020 AAAACCCCACTCAACCATCCAAACCAATCTAATACAGTCCGATGCTGGTGGGTCGAGCGG 1079
Qy 347 eThrThrValPheThrGlyGlnCysPheIleAspArgAsnGlyLysGluValLeuLysTrp 367
Db 1080 CCCTCACCAAGGACAGAGTGGAGATTTTTCACGAAGGCCAGTGGGGTACGGTGTGTGACGA 1139

Qy 367 hrMetTrpLeuLeuArgSerSerVal-AsnAspIleGlyAsp----AspTrpLysAlaThr 385
Db 1140 CCGC-TGGGAACCTGCGTGGAGGACTGCTGCTGCAGAGAGCTTGGGATACAAAGGTGTT 1198
Qy 386 ArgValGlyIleAsn-----IlePheThrArgLeuArgThrGln 398
Db 1199 AAGGTGTGATAGCGAGCTTATTTTGGAAAAAGGTACGGGTCCAA 1243
RESULT 2
5510466-1
; Patent No. 5510466
; APPLICANT: KREIGER, MONTY;KODAMA, TATSUHIKO
; TITLE OF INVENTION: SCAVENGER RECEPTOR PROTEIN AND ANTIBODY
; THEREETO
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,400
; FILING DATE: 16-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 997,113
; FILING DATE: 24-DEC-1992
; APPLICATION NUMBER: 391,486
; FILING DATE: 09-AUG-1989
; APPLICATION NUMBER: 272,002
; FILING DATE: 15-NOV-1988
; SEQ ID NO:1:
; LENGTH: 1807
5510466-1
Alignment Scores:
Pred. No.: 5,76-154 Length: 1807
Score: 1381.50 Matches: 288
Percent Similarity: 85.2% Conservative: 16
Best Local Similarity: 80.7% Mismatches: 34
Query Match: 67.1% Indels: 20
DB: 9 Gaps: 4
US-10-618-570-2 (1-400) x 5510466-1 (1-1807)
Qy 1 MetAlaGlnTrpAspAspPheProAspGlnGlnAspThrAspSerCythrGluSer 20
Db 36 ATGGCACAGTGGGATGACTTTCTCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCT 95
Qy 21 VallysPheAspAlaAatqSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 96 GTGAAGTTCGATGCTCGCTCAGTACAGACTTTGCTCTCCCTCAATCTTAAATAATGGCCCA 155
Qy 41 ThrLeuGlnGluArgMetLysSerTyrlsThrAlaLeuLeuThrLeuTyrlsVal 60
Db 156 ACTCTCAAGAGAGTGAAGTCTTATAAAGTGCACCTGATCACCCTTTATCTCATTTG 215
Qy 61 PheValValLeuValProIlelleGlyIleValAlaAlaGlnLeuLeuTyrlsGluThr 80
Db 216 TTTGTAGTTCCTGTCCTCCCATCTTGGCATAGTGGCAGCTCAGCTCTCCCTCAATGGGAAACG 275
Qy 81 LysAsnCythrValGlySerValAsnAlaAspIleSerProSerProGluGlyGly 100
Db 276 AAGAATTGCACGGTTGGCTCAGTTAATGCAGATATATCTCCAGTCCGGAAGGCAAGGA 335
Qy 101 AsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsnMet 120
Db 336 AATGGCAGTGAAGATGAATGAGATTTCCAGAGCTGTGATGGAAAGCATGAGCAACATG 395
Qy 121 GluSerArgIleGlnTyLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsnPhe 140
Db 396 GAAAGCAATCCAGTATCTTTCCAGATTAATGAAGCCAAATCTCTAGATGCTAAGAAATTC 455
Qy 141 GlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsnSer 160
Db 456 CAAAATTTCCAGCATTAACCACTGATCAAGATTTAATGATGTTCTTTTCCAGCTAAATTC 515
Qy 161 LeuLeuSerSerIleGlnGluHisGluAsnIlelleGlyAspIleSerLysSerLeuVal 180

Db 516 TTACTTCTCCATCCAGAACATGAGAAATATCATAGGGGATATCTCCAGTCAATAGTA 575
Qy GlyLeuAenThrThrValLeuAenGlnPheSerIleGluThrLeuAenGlyArgVal 200
Db 576 GGTCTGAACACACACAGTACTTGAATTCAGTTCAGTATTTGAACACTGAATGGCAGATC 635
Qy GlnGluAenAlaPheLeysGlnGlnGluMetArgLysLeuGluGluArgIleTyrAen 220
Db 636 CAAGAGAAATGATTTAAACAAACAGAGAGATGGTAAATAGAGGCGGTATATACAA 695
Qy AlaSerAlaGluIleLeysSerLeuAenGlnValTyrLeuGluGlnGluLeys 240
Db 696 GCATCAGCAGAAATTAAGTCTCTAGATGAAACAAAGTATATTTGGAAACAGAAATATA 755
Qy GlyGluMetLysLeuLeuAenAenIleThrAenAenLeuAenGluAenThrIleAen 260
Db 756 GGGGAAATGAATGTTTGAATATATCTAATGATCTGAGCTGAAGGATGGGACAT 815
Qy 261 SerGlnThrLeuLysAenIleThrLeuGlnGlnGlyAla----- 273
Db 816 TCTCAGACATTTGAAATATATCTTACTTCAAGGTCTCTCGGACCTCCAGGTGCAAAA 875
Qy 274 -----ArgLysCysSerLeuThrGlyLysThrAenAenLeuGly---SerAenMet 290
Db 876 GGAGATAGAGGCC-TCTCGACA---AAATGGTATACAGGCTTCCAGGTCTAATAGG 931
Qy 291 ThrIleGlyAlaValAenSerArgGlyGluPheThrGlyThrIleThrAlaValThr 310
Db 932 TACTCCAGGTCTTAAAGGTGATCGGGGATCTCTGTTTACTCGGATTCGAGGATCCCC 991
Qy 311 AlaThrSerAenGluIleLeysGluSerProLeuHisGlyThrGlnAenThrIleAenLys 330
Db 992 AGGACCAATGGGAAGACCGGAAGCCAGGACTTATTGGACAAAAAGCCAGAGGAGA 1051
Qy 331 ArgThrGlnProThrPheGlyPheThrValAenThrLysPheSerGluSer 347
Db 1052 AAAAGGGAG-----TGAAGCATGCAAGAGACC 1078

RESULT 3

US-09-949-016-3387
; Sequence 3387, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED.
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3387
; LENGTH: 1366
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3387

Alignment Scores:
Pred. No.: 7,95e-114 Length: 1366
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 3 Gaps: 1

US-10-618-570-2 (1-400) x US-09-949-016-3387 (1-1366)

Qy 1 MetAlaGlnThrAspAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 67 ATGGAGCAGTGGGATCACTTTTCAATCAACAGAGGACACTGATAGTGTCTCCGAATCT 126
Qy 21 ValLysPheAspAlaAArgSerValThrAlaLeuLeuProProHisProLysAenGlyPro 40
Db 127 GTGAATTTGATGCTCGCTCAATGACAGCTTTGCTCTCCGATCTCTTAAACAGACCT 186
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuLysLeuVal 60
Db 187 TCCCTTCAAGAGAAAGTGAAGTCTTCAAGCTGCACCTGATTCCTTTACCTCTCGTG 246
Qy 61 PheValValLeuValProIleGlyLeValAlaGlnLeuLeuLysTrpGluThr 80
Db 247 TTTGAGTTCTCATCCCTCTCATTTGAATAGTGGCAGCTCAACTCTCTGAAGTGGGAACG 306
Qy 81 LysAenCysThrValGlySerValAenAla---AspIleSerProSerProGluGlyLys 99
Db 307 AAGAAATGCTCAGTTAGTTCAACTAATGCAATGATATACTCAAGTCTCACGGGAAAA 366
Qy 100 GlyAenGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAen 119
Db 367 GGAATGACAGGAGAGGAATGAGATTTCAAGAGCTTTTATGGAACACATGAGCAAC 426
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAenGluAlaAenLeuLeuAenAlaLysAen 139
Db 427 ATGGAAGAAGAGATCCAGCATATTTAGACATGGAAGCCAACTCATGACACAGAGCAT 486
Qy 140 PheGlnAenPheSerIleThrAspGlnArgPheAenAspValLeuPheGlnLeuAen 159
Db 487 TTCCAAATTTTCAAGATGACAACTGATCAAGATTTAATGACATTTCTCTCGACGTAAGT 546
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAenIleIleGlyAspIleSerLysSerLeu 179
Db 547 ACCTTGTTTCTCAGTCCAGGACATGGGATGCAATGATGATGAATCTCCAGTCTTAC 606
Qy 180 ValGlyLeuAenThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAenGlyArg 199
Db 607 ATAAGTTTGAATACCATCTGCTTGAATTTGAGCTCAACATAGAAAATCTGAATGGCAA 666
Qy 200 ValGlnGluAenAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 667 ATCCAGAGAAATACCTTCAACAAACAGAGAAATCAGTAAATTAGAGAGCGTGTTCAC 726
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
Db 727 AATGATATCAGCAGAAATATTGCGCTATGAAGAGAAACAAAGTGCATTTGGAAACAGGAATA 786
Qy 240 LysGlyGluMetLysLeuLeuAenAenIleThrAenAenLeuAenGluAenTrpGlu 259
Db 787 AAGGAGAGAGTGAAGTACTGAATAACATCACTAATGATCTCAGACTGAAAGATTGGCAA 846
Qy 260 HisSerGlnThrLeuLysAenIleThrLeuGlnGly 272
Db 847 CATCTCAGACCTTGAAGAAATATCACTTTAATCAAGT 885

RESULT 4

US-08-453-117-3
; Sequence 3, Application US/08453117
; Patent No. 5683903
; GENERAL INFORMATION:
; APPLICANT: Lysko, Paul G.
; APPLICANT: Elshourbagy, Nabil A.
; APPLICANT: Branner, Mary E.
; TITLE OF INVENTION: Attachment Enhanced 293 Cells
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithline Beecham - Corporate Patents
; ADDRESS: U.S.
; STREET: Mailcode - UW2220, 709 Swedeland Road
; CITY: King of Prussia
; STATE: Pennsylvania

Pred. No.: 7,96e-114 Length: 1367
 Score: 1042.50 Matches: 197
 Percent Similarity: 87.2% Conservative: 41
 Best Local Similarity: 72.2% Mismatches: 34
 Query Match: 50.6% Indels: 1
 DB: 2 Gaps: 1

US-10-618-570-2 (1-400) x US-08-948-222-3 (1-1367)

Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
 Db 67 ATGGAGCAGTGGGATCACTTTTACAAATCAACAGAGGACACTGATAGCTGCCGAATCT 126
 Qy 21 VallysPheAspAlaAArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
 Db 127 GTGAATTTGATGCTCGCTCAATGACAGCTTGTCTCCGAATCTCTAAAGACAGCCT 186
 Qy 41 ThrLeuGlnGluArgMetLysSerTyrrLysThrAlaLeuLeuThrLeuLeuVal 60
 Db 187 TCCCTTCAGAGAACTGAAGTCTTCNAAGCTGACAGTATGCTTACCTCTCGTG 246
 Qy 61 PheValValLeuValProIleGlyLeuValAlaAlaGlnLeuLeuLysTrpGluThr 80
 Db 247 TTTGAGTTCATCTCTCTCATTTGGAATAGTGCAGCTCACTCACTTGAAGTGGAAACG 306
 Qy 81 LysAsnCysThrValGlySerValAsnAla---AspLysSerProSerProGluGlyLys 99
 Db 307 AAGAATTTGCTCAGTTAGTTCACTAATGCAATGATATCACTCAAGTCTCACGGGAAA 366
 Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
 Db 367 GGAATGACAGCGAAGGAATGAGATTTCAAGAGTCTTTATGGAACACATGAGCAAC 426
 Qy 120 MetGluSerArgIleGlnTyrrLysSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
 Db 427 ATGGAGAGAAATCCAGCATATTTAGACATGAGGCACTCATGACACAGAGCAT 486
 Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
 Db 487 TTCCAAATTTTCAAGTACCACTGATCAAAAGATTTAATGACATTTCTTCGACGTAAGT 546
 Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
 Db 547 ACCTTGTCTTCTCAGTCCAGGACATGGGAATGCAATAGATGAATCTCCAACTCTCTTA 606
 Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
 Db 607 ATAAGTTTGAATACCATCTCTGATTTGCGCTCACTCAATAGAAATCTGATGGCAAA 666
 Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyrr 219
 Db 667 ATCCAAGAGAAATACCTTCAAAACAACAGAGAAATAGTAAATTTAGAGGAGCGTGTAC 726
 Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrrLeuGluGlnGluIle 239
 Db 727 AATGATATCAGCAGAAATATGCTGATGAAAGAACAAAGTGCATTTTGGAAACAGAAATA 786
 Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
 Db 787 AAGGAGAGTGAAGTACTGTAATACATCACTAATGATCTCAGACTGAAGATGGAA 846
 Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
 Db 847 CATCTCAGACCTTTGAGAAATATCACTTTAATTCAGGT 885

RESULT 6

US-08-973-145-3
 ; Sequence 3, Application US/08973145
 ; Patent No. 5919636
 ; GENERAL INFORMATION:
 ; APPLICANT: Lysko, Paul G.
 ; APPLICANT: Eishourbagy, Nabil A.
 ; APPLICANT: Brawner, Mary E.

TITLE OF INVENTION: Attachment Enhanced 293 Cells
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SmithKline Beecham - Corporate Patents U.S.
 STREET: Mailcode - UW2220, 709 Swedeland Road
 CITY: King of Prussia
 STATE: Pennsylvania
 COUNTRY: U.S.A.
 ZIP: 19406-5090
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/973,145
 FILING DATE: 26-NOV-1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Elizabeth J. Hecht
 REGISTRATION NUMBER: P-41, 824
 REFERENCE/DOCKET NUMBER: P50338
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (610) 270-5009
 TELEFAX: (610) 270-5090
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1367 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: not relevant
 MOLECULE TYPE: cDNA to mRNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 67...1143
 US-08-973-145-3

Alignment Scores:
 Pred. No.: 7,96e-114 Length: 1367
 Score: 1042.50 Matches: 197
 Percent Similarity: 87.2% Conservative: 41
 Best Local Similarity: 72.2% Mismatches: 34
 Query Match: 50.6% Indels: 1
 DB: 2 Gaps: 1

US-10-618-570-2 (1-400) x US-08-973-145-3 (1-1367)

Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
 Db 67 ATGGAGCAGTGGGATCACTTTTACAAATCAACAGAGGACACTGATAGCTGCCGAATCT 126
 Qy 21 VallysPheAspAlaAArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
 Db 127 GTGAATTTGATGCTCGCTCAATGACAGCTTGTCTCCGAATCTCTAAAGACAGCCT 186
 Qy 41 ThrLeuGlnGluArgMetLysSerTyrrLysThrAlaLeuLeuThrLeuLeuVal 60
 Db 187 TCCCTTCAGAGAACTGAAGTCTTCNAAGCTGACAGTATGCTTACCTCTCGTG 246
 Qy 61 PheValValLeuValProIleGlyLeuValAlaAlaGlnLeuLeuLysTrpGluThr 80
 Db 247 TTTGAGTTCATCTCTCTCATTTGGAATAGTGCAGCTCACTCACTTGAAGTGGAAACG 306
 Qy 81 LysAsnCysThrValGlySerValAsnAla---AspLysSerProSerProGluGlyLys 99
 Db 307 AAGAATTTGCTCAGTTAGTTCACTAATGCAATGATATCACTCAAGTCTCACGGGAAA 366
 Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
 Db 367 GGAATGACAGCGAAGGAATGAGATTTCAAGAGTCTTTATGGAACACATGAGCAAC 426
 Qy 120 MetGluSerArgIleGlnTyrrLysSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139

Db 427 ATGGAGAGAGAAATCCAGCATATTTAGACATGGAAGCCAACTCATGGACACAGAGCAT 486
Qy 140 PheGlnAenPheSerIleThrThrAspGlnArgPheAenAspValLeuPheGlnLeuAen 159
Db 487 TTCCAAAATTTAGCATGACAACTGATCAAGATTTAATGACATTTCTTCGAGCTAAGT 546
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAenIleGlyAspIleSerLysSerLeu 179
Db 547 ACCTTGTTTCTCAGTCAGGCAGACATGGGAATGCATAGATGAATTCCTCAAGTCCTTA 606
Qy 180 ValGlyLeuAenThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAenGlyArg 199
Db 607 ATAAGTTTGAATACCACTGCTGATTTGAGCTCAACATAGAAAATCTGAATGGCAAA 666
Qy 200 ValGlnGluAenAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 667 ATCCAAAGAGATACCTTCAACAAACAGAGAAATACAGTAAATTAGAGGCGGTGTTCAC 726
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
Db 727 AATGTATCAGCAGAAATTTATGCTATGAAGAGACAGATGCTAATTTGGACAGGAATA 786
Qy 240 LysGlyGluMetLysLeuLeuAenAenIleThrAenAspLeuArgLeuLysAspTyrGlu 259
Db 787 AAAGGAGAGTGAAGTACTGAATAACATCACTAATGATCTCAGACTGAAAGATTGGAA 846
Qy 260 HisSerGlnThrLeuLysAenIleThrLeuLeuGlnGly 272
Db 847 CATTTCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 885

RESULT 7

US-09-016-434-1260
; Sequence 1260, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016.434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1260:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1367 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; IMMEDIATE SOURCE:
; LIBRARY: GENEANK
; CLONE: g219991
US-09-016-434-1260
Alignment Scores:
Pred. No.: 7.96e-114 Length: 1367
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 3 Gaps: 1

US-10-618-570-2 (1-400) x US-09-016-434-1260 (1-1367)

Qy 1 MetAlaGlnTrpAspAspPheProAspGlnGlnAspThrAspSerCysThrGluSer 20
Db 67 ATGGAGCAGTGGGATCACTTTTCAATCAACAGAGGACACTGATAGCTGCTCCGAATCT 126
Qy 21 ValIysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAenGlyPro 40
Db 127 GTGAATTTGATGCTCGCTCAATGACAGCTTTGCTTCTCGAATCTTAAACACAGCCCT 186
Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60
Db 187 TCCCTTCAAGAGAACTGAACTGCTTCAAGCTGCACTGATTCCTTACCTCTCGTG 246
Qy 61 PheValValLeuValProIleGlyIleValAlaAlaGlnLeuLeuLysTyrGluThr 80
Db 247 TTTGCACTTCTCATCCCTCTCATTTGGAATAGTGCAGCTCAACTCTCGAATGCGGAAACG 306
Qy 81 LysAenCysThrValGlySerValAenAla---AspIleSerProSerProGluGlyLys 99
Db 307 AAGAAATGCTCAGTTAGTTCAACTTAATGCAATATATACTCAAGTCTCACGGGAAA 366
Qy 100 GlyAenGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAen 119
Db 367 GGAATACACAGCAGAGAGAAATGAGATTTCAAGAAGCTTTATGGAACACATCAGCAAC 426
Qy 120 MetGluSerArgIleGlnTyrLeuSerAspAenGluAlaAenLeuLeuAspAlaLysAen 139
Db 427 ATGGAGAGAGAAATCCAGCATATTTTATAGACATGGAAGCCCAACTCATGGACACAGAGCAT 486
Qy 140 PheGlnAenPheSerIleThrThrAspGlnArgPheAenAspValLeuPheGlnLeuAen 159
Db 487 TTCCAAAATTTAGCATGACAACTGATCAAGATTTAATGACATTTCTTCGAGCTAAGT 546
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAenIleIleGlyAspIleSerLysSerLeu 179
Db 547 ACCTTGTTTCTCAGTCAGGCAGACATGGGAATGCATAGATGAATTCCTCAAGTCCTTA 606
Qy 180 ValGlyLeuAenThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAenGlyArg 199
Db 607 ATAAGTTTGAATACCACTGCTGATTTGAGCTCAACATAGAAAATCTGAATGGCAAA 666
Qy 200 ValGlnGluAenAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
Db 667 ATCCAAAGAGATACCTTCAACAAACAGAGAAATACAGTAAATTAGAGGCGGTGTTCAC 726
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
Db 727 AATGTATCAGCAGAAATTTATGCTATGAAGAGACAGATGCTAATTTGGACAGGAATA 786
Qy 240 LysGlyGluMetLysLeuLeuAenAenIleThrAenAspLeuArgLeuLysAspTyrGlu 259
Db 787 AAAGGAGAGTGAAGTACTGAATAACATCACTAATGATCTCAGACTGAAAGATTGGAA 846
Qy 260 HisSerGlnThrLeuLysAenIleThrLeuLeuGlnGly 272
Db 847 CATTTCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 885

RESULT 8

PCT-US96-08081-3

```

; Sequence 3, Application PC/TUS9608081
; GENERAL INFORMATION:
; APPLICANT: SmithKline Beecham Corporation
; TITLE OF INVENTION: Attachment Enhanced 293 Cells
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham - Corporate Patents U.S.
; STREET: Mailcode - UW2220, 709 Swedeland Road
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19406-5090
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/08081
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jervis, Herbert H.
; REGISTRATION NUMBER: 31,171
; REFERENCE/DOCKET NUMBER: P50338
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 270-5019
; TELEFAX: (610) 270-5090
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1367 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 67..1143
; PCT-US96-08081-3

Alignment Scores:
Pred. No.: 7,96e-114 Length: 1367
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 6 Gaps: 1

US-10-618-570-2 (1-400) x PCT-US96-08081-3 (1-1367)
Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 67 ATGGAGCAGTGGGATCACTTTCACAAATCAACAGAGGACACTGATAGTCTCGAATCT 126
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 127 GTGAAATTGATGCTCGCTCAATGACAGCTTGTCTTCGGAATCTTAAACACAGCCCT 186
Qy 41 ThrLeuGlnGluArgMetLysSerTyrlsThrAlaLeuLeuThrLeuTyrlsVal 60
Db 187 TCCCTTCAAGAGAAAGTGAAGTCTTCAAGCTGCACCTGATTCCTTACCTCTCGTG 246
Qy 61 PheValValLeuValProIleGlyLeValAlaAlaGlnLeuLeuLysTrpGluThr 80
Db 247 TTTGCGATTCTATCCCTCTCTCTGATGATGAGTGCAGCTCACTCTCGAAGTGGAAACG 306
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 307 AAGAATTGCTCAGTTAGTTCAACTAATGCAATGATATCACTCAAGTCTCACGGGAAA 366
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 367 GGAATGACAGCGAAGAGGAATGAGATTTCAAGAGTCTTTATGGAACACATGAGCAAC 426

```

```

Qy 120 MetGluSerArgIleGlnIleTyrlsLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 427 ATGGAGAAGAGAAATCCAGCATATTTAGACATGATGAAGCCCACTCATGGACACAGAGCAT 486
Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 487 TTCCAAATTTTCAGCATGACAACTGATCAAGATTTTAATGACATCTTCTCGACGTAAGT 546
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
Db 547 ACCTTGTCTTCTCAGTCCAGGACATGGGAATGCAATAGATGAATCTCCAAGTCTCTTA 606
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 607 ATAGTTTGAATACCACTATGCTTGAATTCAGCTCAACATAGAAAATCTGAATGGCAAA 666
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluArgIleTyrls 219
Db 667 ATCCAGAGAAATACCTTCAACAAACAAAGAGAAATCAGTAAATTAGAGGCGGTGTTTAC 726
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrlsLeuGluGlnGluIle 239
Db 727 AATGATCAGCAGAAATATGCTGCTATGAAGAGAACAGTGCATTTGGACACAGAAATA 786
Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
Db 787 AAAGGAGAACTGAAAGTACTGAATAACATCACTAATGATCTCAGACTGAAAGATTGGGA 846
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 847 CATTCAGACCTTGAGAAATATCACTTTAATCAAGGT 885

```

RESULT 9

```

US-08-453-117-1
; Sequence 1, Application US/08453117
; Patent No. 5683903
; GENERAL INFORMATION:
; APPLICANT: Lysko, Paul G.
; APPLICANT: Elehourbagy, Nabil A.
; APPLICANT: Brawner, Mary E.
; TITLE OF INVENTION: Attachment Enhanced 293 Cells
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham - Corporate Patents
; ADDRESSEE: U.S.
; STREET: Mailcode - UW2220, 709 Swedeland Road
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19406-5090
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,117
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jervis, Herbert H.
; REGISTRATION NUMBER: 31,171
; REFERENCE/DOCKET NUMBER: SBC-P50338
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 270-5019
; TELEFAX: (610) 270-5090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2028 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant

```

```

; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 47..1402
US-08-453-117-1
Alignment Scores:
Pred. No.: 1,5e-113 Length: 2028
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservativeness: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 2 Gaps: 1

US-10-618-570-2 (1-400) x US-08-453-117-1 (1-2028)
Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 47 ATGGAGCAGTGGGATCACTTTTCAACAATCAACAGGAGGACACTGATAGCTGCTCCGAATCT 106
Qy 21 VallysPheAspAlaAatqSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 107 GTGAAATTTGATGCTCGCTCAATGACAGCTTGTCTTCCGAATCTTAAATAACAGCCCT 166
Qy 41 ThrLeuGlnGluArgMetLysSerTyrlsThrAlaLeuLeuThrLeuTyrlsLeuVal 60
Db 167 TCCCTTCAAGAGAACTGAGTCTTCAAAAGCTGCACCTGATTTGCCCTTACTCTCGTG 226
Qy 61 PheValValLeuValProIleIleGlyLeValAlaAlaGlnLeuLeuTyrlsGluThr 80
Db 227 TTTCAGTTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTGAAAGTGGAAACG 286
Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 287 AAGAAATTCCTCAGTTAGTTAGTCACTAATGCAATGATATACTCAAAAGTCTCAGCGGAAA 346
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 347 GGAATGACAGGAGAGGAATGAGATTTCAAGNAGCTTTTATGGAACACATGAGCAAC 406
Qy 120 MetGluSerArgIleGlnTyrlsLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 407 ATGGAGAGAGAAATCCAGCATATTTAGACATGGAAGCAACCTCATGGACACAGAGCAT 466
Qy 140 PheGlnAsnPheSerIleThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 467 TTCCAAAATTTTCAGCATGACCACTGATCAAAAGATTTAATGACATTTCTTCGACGTAA 526
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
Db 527 ACCTTGTTTCTCAGTCCAGGACATGGGAATGCAATGATGATGAATCTCCAAATCTCTTA 586
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 587 ATAAGTTTGAATACCAATGCTTGTGTTGAGCTCAACATAGAAAATCTGAAATGGCAAA 646
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGlnGluMetArgLysLeuGluGluArgIleTy 219
Db 647 ATCCAAGAGAAATACCTTCAAAACAACAGAGAAATACAGTAAATTTAGAGGACGCTGTTTAC 706
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrlsLeuGluGlnGluIle 239
Db 707 AATGATCAGCAGAAATTTATGGCTATGAAAGAGACACAGTGCATTTTGGAAACAGGAATA 766
Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
Db 767 AAAGAGAGAGTGAAGTACTGAATAACATCACTAATGATCTCAGACTGAAAGATTGGAA 826
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 827 CATTTCTCAGACCTTGAGAAATATCACTTTTAATTCAGGT 865
RESULT 10
```

```

US-08-948-222-1
; Sequence 1, Application US/08948222
; Patent No. 5863798
; GENERAL INFORMATION:
; APPLICANT: Lysko, Paul G.
; APPLICANT: Elshourbagy, Nabil A.
; APPLICANT: Brawner, Mary E.
; TITLE OF INVENTION: Attachment Enhanced 293 Cells
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham - Corporate Patents
; ADDRESSEE: U.S.
; STREET: Mailcode - UW2220, 709 Swedeland Road
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19406-5090
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,222
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,117
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jervis, Herbert H.
; REGISTRATION NUMBER: 31,171
; REFERENCE/DOCKET NUMBER: SBC-P50338
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 270-5019
; TELEFAX: (610) 270-5090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2028 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: No. 5863798 Relevant
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 47..1402
US-08-948-222-1
Alignment Scores:
Pred. No.: 1,5e-113 Length: 2028
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservativeness: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 2 Gaps: 1

US-10-618-570-2 (1-400) x US-08-948-222-1 (1-2028)
Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 47 ATGGAGCAGTGGGATCACTTTTCAACAATCAACAGGAGGACACTGATAGCTGCTCCGAATCT 106
Qy 21 VallysPheAspAlaAatqSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 107 GTGAAATTTGATGCTCGCTCAATGACAGCTTGTCTTCCGAATCTTAAATAACAGCCCT 166
Qy 41 ThrLeuGlnGluArgMetLysSerTyrlsThrAlaLeuLeuThrLeuTyrlsLeuVal 60
Db 167 TCCCTTCAAGAGAACTGAGTCTTCAAAAGCTGCACCTGATTTGCCCTTACTCTCGTG 226
Qy 61 PheValValLeuValProIleIleGlyLeValAlaAlaGlnLeuLeuTyrlsGluThr 80
Db 227 TTTCAGTTCATCCCTCTCATTTGGAATAGTGGCAGCTCAACTCTGAAAGTGGAAACG 286
```

QY 81 LysAsnGlyThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
 DB 287 AAGAAATTCCTCAGTTAGTTCAACTAATCAATGATATAACTCAAGGTCTCACGGGAAA 346
 QY 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
 DB 347 GGAATGACAGGAGAGGAATGATTTCAAGAGTCTTTATGGAACACATGAGCAAC 406
 QY 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
 DB 407 ATGGAAGAAGAAATCCAGCATATTGACATGGAAGCCAACTCATGGACACAGAGCAT 466
 QY 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
 DB 467 TTCAAAATTTTCAGCATGACAACTGATCAAAAGATTTAATGACATTTCTTCGACGTAAGT 526
 QY 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
 DB 527 ACCTTGTTTCTCAGTCCAGGACATGGGAATGCAATGATGAATCTCCAGTCCCTTA 586
 QY 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
 DB 587 ATAAGTTTGAATACCATCTGTTGATTTGCGAGCTCAACATAGAAAATCTCAATGGCAAA 646
 QY 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
 DB 647 ATCCAAAGAGATACCTTCAACACAGAGGAAATCAGTAAATTAGAGGACGCTGTTTAC 706
 QY 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
 DB 707 AATGATCAGCAGAAATTTATGGCTATGAAAGAGAACAAAGTGCAATTTGGAACAGGAATA 766
 QY 240 LysGlyGluMetLysLeuAsnAsnIleThrLeuLeuGlnGly 272
 DB 827 CATCTCAGACCTTGAGAAATATCATCTTATTTCAAGT 865

RESULT 11

US-08-973-145-1
 ; Sequence 1, Application US/08973145
 ; Patent No. 5919636
 ; GENERAL INFORMATION:
 ; APPLICANT: Lysko, Paul G.
 ; APPLICANT: Elshourbagy, Nabil A.
 ; APPLICANT: Brawer, Mary E.
 ; TITLE OF INVENTION: Attachment Enhanced 293 Cells
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SmithKline Beecham - Corporate Patents U.S.
 ; STREET: Mailcode - UW2220, 709 Swedeland Road
 ; CITY: King of Prussia
 ; STATE: Pennsylvania
 ; COUNTRY: U.S.A.
 ; ZIP: 19406-5090
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/973.145
 ; FILING DATE: 26-NOV-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Elizabeth J. Hecht
 ; REGISTRATION NUMBER: P-41, 824
 ; REFERENCE/DOCKET NUMBER: P50338
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (610) 270-5009

; TELFAX: (610) 270-5090
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2028 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: not relevant
 ; MOLECULE TYPE: cDNA to mRNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 47..1402
 ; US-08-973-145-1

Alignment Scores:
 Pred. No.: 1.5e-113 Length: 2028
 Score: 1042.50 Matches: 197
 Percent Similarity: 87.2% Conservative: 41
 Best Local Similarity: 72.2% Mismatches: 34
 Query Match: 50.6% Indels: 1
 DB: 2 Gaps: 1

US-10-618-570-2 (1-400) x US-08-973-145-1 (1-2028)

QY 1 MetAlaGlnTyrAspAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
 DB 47 ATGAGCAGTGGGATCACTTTCAATCAACAGAGGACACTGATAGCTCTCGAATCT 106
 QY 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProProHisProLysAsnGlyPro 40
 DB 107 GTGAATTTGATGCTCCTCAATGACAGCTTTGCTTCGATCTCTAATAACAGCCCT 166
 QY 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuIleThrLeuTyrLeuIleVal 60
 DB 167 TCCCTTCAAGAGAACTGAAGTCTTCAAGCTGCACCTGATTGGCTTTACCTCTCGTG 226
 QY 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTyrGluThr 80
 DB 227 TTTGCAGTTCTCATCTCTCTCATTTGGAATAGTGGCAGCTCAACTCTGAACTGGGAAACG 286
 QY 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
 DB 287 AAGAAATTCCTCAGTTAGTTCAACTAATGCAATGATATACTCAAGTCTCACGGGAAA 346
 QY 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
 DB 347 GGAATGACAGGAGAGGAATCAGATTTCAAGAGTCTTTATGGAACACATGAGCAAC 406
 QY 120 MetGluSerArgIleGlnTyrLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
 DB 407 ATGGAAGAAGAGATCCAGCATATTTTAGACATGGAAGCCAACTCATGGACACAGAGCAT 466
 QY 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
 DB 467 TTCAAAATTTTCAGCATGACAACTGATCAAAAGATTTAATGACATTTCTTCAGCTAAGT 526
 QY 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
 DB 527 ACCTTGTTTCTCAGTCCAGGACATGGGAATGCAATGATGAATCTCCAGTCCCTTA 586
 QY 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
 DB 587 ATAAGTTTGAATACCATCTGTTGATTTGCGAGCTCAACATAGAAAATCTCAATGGCAAA 646
 QY 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
 DB 647 ATCCAAAGAGATACCTTCAACACAGAGGAAATCAGTAAATTAGAGGACGCTGTTTAC 706
 QY 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
 DB 707 AATGATCAGCAGAAATTTATGGCTATGAAAGAGAACAAAGTGCAATTTGGAACAGGAATA 766
 QY 240 LysGlyGluMetLysLeuAsnAsnIleThrLeuLeuGlnGly 259

Db 767 AAGGAGAGTGAAGTACTGAATAACATCACTAATGATCTCAGACTGAAGATTGGGAA 826
PCT-US96-08081-1
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 827 CATTCTCAGACCTTGAGAAATATCATCTTAATCAAGGT 865
RESULT 12
; Sequence 1, Application PC/TUS9608081
; GENERAL INFORMATION:
; APPLICANT: SmithKline Beecham Corporation
; TITLE OF INVENTION: Attachment Enhanced 293 Cells
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham - Corporate Patents U.S.
; STREET: Mailcode - UW2220, 709 Swedeland Road
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19406-5090
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/08081
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jervis, Herbert H.
; REGISTRATION NUMBER: 31,171
; REFERENCE/DOCKET NUMBER: P50338
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 270-5019
; TELEFAX: (610) 270-5090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2028 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 47..1402
PCT-US96-08081-1
Alignment Scores:
Pred. No.: 1.5e-113 Length: 2028
Score: 1042.50 Matches: 197
Percent Similarity: 87.2% Conservative: 41
Best Local Similarity: 72.2% Mismatches: 34
Query Match: 50.6% Indels: 1
DB: 6 Gaps: 1
US-10-618-570-2 (1-400) x PCT-US96-08081-1 (1-2028)
Qy 1 MetAlaGlnTrpAspPheProAspGlnGlnGluAspThrAspSerCysThrGluSer 20
Db 47 ATGGAGCAGTGGGATCACTTTTCACAAATCAACAGGAGGACACTGATAGCTGCTCGAATCT 106
Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
Db 107 GTGAAATTTGATGCTCGCTCAATGACAGCTTTGCTTCGGAATCTCTAAACACAGCCCT 166
Qy 41 ThrLeuGlnGluArgMetLysSerTyrlsThrAlaLeuIleThrLeuTyrlleVal 60
Db 167 TCCCTTCAAGAGAACTCAAGTCTTCAAGAGCTGCAGCTGATGCTTTCCTCTCTCGTG 226
Qy 61 PheValValLeuValProIleIleGlyIleValAlaAlaGlnLeuLeuLysTrpGluThr 80
Db 227 TTTCAGTCTTCATCCCTCTCATTTGGATAGTGGCAGCTCACTCTCACTGAAAGTGGGAAACG 286

Qy 81 LysAsnCysThrValGlySerValAsnAla---AspIleSerProSerProGluGlyLys 99
Db 287 AAGAAATTCCTAGTTAGTTCACTAATGCAATGATATNACTCAAGTCTCACCAGGAAA 346
Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
Db 347 GGAATTCACAGCGAAGAGGAAATGAGATTTCAAGAAAGTCTTTATGGAACACATGAGCAAC 406
Qy 120 MetGluSerArgIleGlnTyrlsLeuSerAspAsnGluAlaAsnLeuLeuAspAlaLysAsn 139
Db 407 ATGGAGAGAGAAATCCAGCATATTTTAGACATGGAAGCCCAACCTCATGGACACAGAGCAT 466
Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
Db 467 TTCAAAATTTTCAGCATGACACATGATCAAGAATTTAATGACATTTCTTCGACGTAAGT 526
Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIleIleGlyAspIleSerLysSerLeu 179
Db 527 ACCTTGTCTTCAGTCCAGGAGACATGGGAATGCAATAGATGAATCTCCAAGTCCCTTA 586
Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
Db 587 ATAAAGTTTGAATACACATATGCTTGTGATTTGCAGCTCAACATAGAAAATCTGAATGGCAAA 646
Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluArgIleTyrl 219
Db 647 ATCCAGAGAAATACCTTCAACAAAGAGAGAAATCAGTAATTTAGAGGACGCTGTTTAC 706
Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrlsLeuGlnGluIle 239
Db 707 AATGATATCAGCAGAAATATGCTATGAAAGAGAACAAAGTGCAATTTTGGACACAGAAATA 766
Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTrpGlu 259
Db 767 AAAGGAGAGTGAAGTACTGAATAACATCACTAATGATCTCAGACTGAAGATTGGGAA 826
Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
Db 827 CATTCTCAGACCTTGAGAAATATCACTTTAATTCAGGT 865
RESULT 13
; Sequence 1, Application US/08154365
; Patent No. 5624904
; GENERAL INFORMATION:
; APPLICANT: Dunne, Dana W.
; APPLICANT: Resnick, David
; APPLICANT: Kreiger, Monty
; APPLICANT: Joiner, Keith A.
; TITLE OF INVENTION: Method for Treating Gram-Positive
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Ga
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/154,365
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: MIT6392

TELECOMMUNICATION INFORMATION:

TELEPHONE: (404)-815-6508
 TELEFAX: (404)-815-6555
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2037 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: homo sapien
 IMMEDIATE SOURCE:
 LIBRARY: THP-1
 PUBLICATION INFORMATION:
 AUTHORS: Ashkenas, et al.
 JOURNAL: J. Lipid Res.
 VOLUME: 34
 PAGES: 983-1000
 DATE: 1993
 RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 2037

US-08-154-365-1

Alignment Scores:
 Pred. No.: 3,44e-113 Length: 2037
 Score: 1039.50 Matches: 196
 Percent Similarity: 87.2% Conservative: 42
 Best Local Similarity: 71.8% Mismatches: 34
 Query Match: 50.5% Indels: 1
 Gaps: 2

US-10-618-570-2 (1-400) x US-08-154-365-1 (1-2037)

Qy 1 MetAlaGlnTrpAspAppPheProAspGlnGluAspThrAspSerCysThrGluSer 20
 Db 47 ATGGAGCAGTGGGATCACTTTTCAATCAATCAACAGGAGGACCTGATAGCTGCTCGGAATCT 106
 Qy 21 ValLysPheAspAlaArgSerValThrAlaLeuLeuProHisProLysAsnGlyPro 40
 Db 107 GTGAATTTGATGTCGCTCAATGACAGCTTTGCTTCGATCTCTAAACACAGCCT 166
 Qy 41 ThrLeuGlnGluArgMetLysSerTyrLysThrAlaLeuLeuThrLeuLysLeuVal 60
 Db 167 TCCCTTCAAGAGAAAGTCAAGTCTTCAAGCTGCACTGATTCCTTACCTTCTCGTG 226
 Qy 61 PheValValLeuValProIleGlyTleValAlaAlaGlnLeuLeuLysTyrGluThr 80
 Db 227 TTTCAGGTTCATCCCTCTCATTTGGAATAGTGCAGCTCAACTCCTGAAGTGGGAACG 286
 Qy 81 LysAsnCysThrValGlySerValAsnAla--AspIleSerProSerProGluGlyLys 99
 Db 287 AAGATTTGCTCAGTTAGTTCAACTAATGCAATATGATTAATCAAGTCTCACGGGAAA 346
 Qy 100 GlyAsnGlySerGluAspGluMetArgPheArgGluAlaValMetGluArgMetSerAsn 119
 Db 347 GGAATGACGCGAAGAGAAATGAGATTTCAAGAGTCTTTATGCAACACATCAGCAAC 406
 Qy 120 MetGluSerArgIleGlnTyrLeuSerAspGlnAlaAsnLeuLeuAspAlaLysAsn 139
 Db 407 ATGGAGAAGAAATCCAGCATATTTAGACATGGAAGCCAACTCATGGACACAGAGCAT 466
 Qy 140 PheGlnAsnPheSerIleThrThrAspGlnArgPheAsnAspValLeuPheGlnLeuAsn 159
 Db 467 TTCCAAATTTTCAGATGACAACTGATCAAGATTTAATGACATTTCTCGCATAGT 526
 Qy 160 SerLeuLeuSerSerIleGlnGluHisGluAsnIlelleGlyAspIleSerLysSerLeu 179
 Db 527 ACCTTGTCTTCTCAGTCCAGGACATGGGAATGCAATAGATGAATCAATCTCAAGTCTTCA 586
 Qy 180 ValGlyLeuAsnThrThrValLeuAspLeuGlnPheSerIleGluThrLeuAsnGlyArg 199
 Db 527 ACCTTGTCTTCTCAGTCCAGGACATGGGAATGCAATAGATGAATCAATCTCAAGTCTTCA 586

Db 587 ATAAGTTTGAATACCACTTGTGATTTCAGCTCAACATAGAAATCTGAATGGCAA 646
 Qy 200 ValGlnGluAsnAlaPheLysGlnGlnGluMetArgLysLeuGluGluArgIleTyr 219
 Db 647 ATCCAAGAGAATACCTTTCAACACAAAGAGAAATCAGTAAATTAGAGGCGGTGTAC 706
 Qy 220 AsnAlaSerAlaGluIleLysSerLeuAspGluLysGlnValTyrLeuGluGlnGluIle 239
 Db 707 AATGTATCAGCAGAAATATTGGCTATGAAAGAAACAAAGTGCATTTTGGAAACAGGAATA 766
 Qy 240 LysGlyGluMetLysLeuLeuAsnAsnIleThrAsnAspLeuArgLeuLysAspTyrGlu 259
 Db 767 AAGAGAGAGTGAAGTACTGATTAACATCACTAATGATCTGACACTGGAAGATTGGNA 826
 Qy 260 HisSerGlnThrLeuLysAsnIleThrLeuLeuGlnGly 272
 Db 827 CATCTCAGACCTTGAGAAATATCATTAAATTCAGGT 865

RESULT 14

US-08-554-586-1
 ; Sequence 1, Application US/08554586
 ; Patent No. 5767379

GENERAL INFORMATION:

APPLICANT: BASZCZYNSKI, Chris
 APPLICANT: HOOD, Elizabeth
 APPLICANT: MADDOCK, Sheila
 APPLICANT: MEYER, Terry Euclyaire
 APPLICANT: REGISTER III, James C.
 APPLICANT: WITCHER, Derrick
 APPLICANT: HOWARD, John A.
 TITLE OF INVENTION: COMMERCIAL PRODUCTION OF AVIDIN IN
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Lardner
 STREET: 3000 K Street, N.W., Suite 500
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20007-5109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/554,586
 FILING DATE: 06-NOV-1995
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 33229/352/PIHI
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)672-5300
 TELEFAX: (202)672-5399
 TELEX: 904136
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 484 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

US-08-554-586-1

Alignment Scores:
 Pred. No.: 2,29e-72 Length: 484
 Score: 688.00 Matches: 130
 Percent Similarity: 98.5% Conservative: 1
 Best Local Similarity: 97.7% Mismatches: 2
 Query Match: 33.4% Indels: 0
 Gaps: 2

US-10-618-570-2 (1-400) x US-08-554-586-1 (1-484)

```
QY 268 ThrLeuLeuGlnGlyAlaArgLysCysSerLeuThrGlyLysThrAsnAspLeuGly 287
Db 69 TCCCTCGCCAGCGCGCCAGGAAGTCTCCCTCACCGCAAGTGACCAATGACTCGGC 128
QY 288 SerAsnMetThrIleGlyAlaValAsnSerArgGlyGluPheThrGlyThrTyrlleThr 307
Db 129 TCCACATGACCATCGCGCGCTGAACTCCAGGGCGGAGTTCCCGGACCTACATCACC 188
QY 308 AlaValThrAlaThrSerAsnGluLeuLysGluSerProLeuHisGlyThrGlnAsnThr 327
Db 189 GCGGTGACCGCGCACCTCCACAGCAGATCAAGAGTCCCCCTCCACGGTACCCAGAACACC 248
QY 328 IleAsnLysArgThrGlnProThrPheGlyPheThrValAsnThrLysPheSerGluSer 347
Db 249 ATCAACAAGAGGACCCAGCCACCTTCGGCTTCACTCCGTAACCTGGAAAGTTCTCCGAGTCC 308
QY 348 ThrThrValPheThrGlyGlnCysPheIleAspArgAsnGlyLysGluValLeuLysThr 367
Db 309 ACCACCGTGTTCACCGCGCATGTCTTCATCCAGCCGCAACGCGCAAGGAGTGCTCAAGACC 368
QY 368 MetTrpLeuLeuArgSerSerValAsnAspIleGlyAspAspTrpLysAlaThrArgVal 387
Db 369 ATGTGGCTCTCGAGGAGCTCCGTGAATGACATCGCGCAGCAGCTGGAAAGCCACCCGCGTG 428
QY 388 GlyIleAsnIlePheThrArgLeuArgThrGlnLysGlu 400
Db 429 GGCATCAACATCTTCAACCGCGCTCCGACCCAGAAAGGAG 467
```

RESULT 15

```
US-08-831-399-3
; Sequence 3, Application US/08831399
; Patent No. 6312916
; GENERAL INFORMATION:
; APPLICANT: Kopetzki, Erhard; Muller, Rainer;
; APPLICANT: Eng, Richard; Schmitt, Urban; Deger, Arno; Brandstetter, Hans
; TITLE OF INVENTION: Recombinant Inactive Core
; TITLE OF INVENTION: Streptavidin Mutants
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2 PC-DOS
; OPERATING SYSTEM:
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/831.399
; FILING DATE: 1-April-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 13 053.0
; FILING DATE: 1-April-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 37 718.8
; FILING DATE: 16-September-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6312916man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 1105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 604 base pairs
; TYPE: nucleic acid
```

```
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 44..499
; OTHER INFORMATION: Positions 44..115 correspond to sig
; OTHER INFORMATION: peptide and 116..499 to mat peptide
US-08-831-399-3
```

```
Alignment Scores:
Pred. No.: 5.07e-71 Length: 604
Score: 678.00 Matches: 128
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 32.9% Indels: 0
DB: 3 Gaps: 0
```

US-10-618-570-2 (1-400) x US-08-831-399-3 (1-604)

```
QY 273 AlaArgLysCysSerLeuThrGlyLysThrAsnAspLeuGlySerAsnMetThrIle 292
Db 116 GCAGAAAGTGCTCGCTGACTGGGAAATGGACCAACGATCTGGGCTCCAACATGACCATC 175
QY 293 GlyAlaValAsnSerArgGlyGluPheThrGlyThrTyrlleThrAlaValThrAlaThr 312
Db 176 GGGGCTGTGAACAGCAGAGGTGAATTCACAGGCACCTACATCACAGCGTAACAGCCACA 235
QY 313 SerAsnGluIleLysGluSerProLeuHisGlyThrGlnAsnThrIleAsnLysArgThr 332
Db 236 TCAAAATGAGATCAAAAGAGTCACCACTGCATGGGACACAAAACACCATCAACAGAGGACC 295
QY 333 GlnProThrPheGlyPheThrValAsnTrpLysPheSerGluSerThrValPheThr 352
Db 296 CAGCCCACTTTGGCTTTCACCGCAATTGGAAGTTTTCAGAGTCCACCATCTTCTCACG 355
QY 353 GlyGlnCysPheIleAspArgAsnGlyLysGluValLeuLysThrMetTrpLeuLeuArg 372
Db 356 GGCAGTGCTTCATAGACAGGAATGGGAAGAGGTCTCTGAAGACCATGTGGCTGTGGGG 415
QY 373 SerSerValAsnAspIleGlyAspAspTrpLysAlaThrArgValGlyIleAsnIlePhe 392
Db 416 TCAAGTGTAAATGACATTTGGTGATGACTGGGAAAGCTACCAGGGTCCGCATCAACATCTTC 475
QY 393 ThrArgLeuArgThrGlnLysGlu 400
Db 476 ACTGCGCTCGCACACAGAAAGGAG 499
```

Search completed: February 27, 2006, 12:42:35
Job time : 239 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2006, 08:22:53 ; Search time 24993 Seconds
(without alignments)
11774.439 Million cell updates/sec

Title: US-10-618-570-1
Perfect score: 5177
Sequence: 1 ttgaagaccaccaccgta.....agcgggggttttcattgg 5177

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

1: gb_pa.*
2: gb_in.*
3: gb_env.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sts.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	5177	100.0	5177	6	BD136828
2	5177	100.0	5177	6	AX018983
3	2078.8	40.2	7699	6	A80474
4	2078.8	40.2	7699	6	BD136838
5	2078.8	40.2	7699	6	AR381629
6	1972.6	38.1	7436	6	C0879093
7	1916.6	37.0	7546	6	BD138566
8	1644.6	31.8	7165	6	BD138564
9	1629.4	31.5	5594	6	AR214682
10	1588.6	30.7	5130	6	AR634978
11	1588.6	30.7	5130	6	AX359935
12	1588.6	30.7	5130	6	AX382149
13	1487.6	28.7	5464	11	SYNMMPLN1
14	1486	28.7	5874	6	BD138565
15	1486	28.7	5874	11	SYNMMPLN4
16	1486	28.7	6141	6	BD138562
17	1486	28.7	6365	6	I70974
18	1486	28.7	6506	13	U00220

19	1486	28.7	6522	6	BD138563
20	1486	28.7	6620	6	AR302094
21	1463	28.3	6501	11	AF132211
22	1455.2	28.1	5865	6	AR164477
23	1452	28.0	6072	11	D88622
24	1448.4	28.0	8591	6	I58322
25	1448.4	28.0	8591	6	I58323
26	1448.4	28.0	8591	6	I60508
27	1448.4	28.0	8591	6	I60509
28	1448.4	28.0	8591	6	I77052
29	1448.4	28.0	8591	6	I77053
30	1448.4	28.0	8591	6	I87173
31	1448.4	28.0	8591	6	I87174
32	1448	28.0	6094	11	AF113968
33	1447.4	28.0	6248	11	AB041928
34	1444.6	27.9	6277	11	AB086385
35	1437.2	27.8	9609	11	AP396261
36	1434	27.7	6700	6	AR585288
37	1434	27.7	8518	6	AR321727
38	1434	27.7	8518	6	AR585290
39	1434	27.7	8518	6	AR656176
40	1429.2	27.6	6700	6	AR321726
41	1429.2	27.6	6700	6	AR656175
42	1425	27.5	5689	6	BD075836
43	1425	27.5	5689	6	BD081937
44	1425	27.5	5689	6	BD269812
45	1425	27.5	5689	6	AR593118

ALIGNMENTS

RESULT 1
BD136828

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BD136828 5177 bp DNA linear PAT 18-SEP-2002
Biotin-binding receptor molecule.
BD136828
BD136828.1 GI:23231773
JP 2002504328-A/1.
unidentified
unclassified.
1 (bases 1 to 5177)
Herttuaia,S.Y., Kulomaa,M., Lehtolainen,P., Marjomaki,V. and Airenne,K.
Biotin-binding receptor molecule
Patent: JP 2002504328-A 1 12-FEB-2002;
ARC THERAPEUTICS LTD
OS Unidentified
PN JP 2002504328-A/1
PD 12-FEB-2002
PF 23-FEB-1999 JP 2000532517
PR 23-FEB-1998 GB 9803757.5,24-JUN-1998 GB 9813653.4 PI
SEPPO YLA HERTTUAILA, MARKKO KULOMAA, PAULIINA LEHTOLAINEN, VARPU PI MARJOMAKI,
PI KARI AIRENNE
PC C12N15/09,A61K48/00,A61P43/00,C07K14/705//A61K38/00,C12N15/00,
PC A61K37/02
CC Strandedness: Single;
CC Topology: linear;
CC Biotin-binding receptor molecule
FH Key Location/Qualifiers
FT CDS 1071..2270.
Location/Qualifiers
1..5177
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

FEATURES
source

ORIGIN

Query Match 100.0%; Score 5177; DB 6; Length 5177;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTTGAAAGACCCCAACCGTAGGTGGCAAGCTAGCTTAAGTAAACGCCACTTTTGCAGGCAT 60
Db |||||
1 TTTGAAAGACCCCAACCGTAGGTGGCAAGCTAGCTTAAGTAAACGCCACTTTTGCAGGCAT 60
Qy 61 GGAATAATACATAATCTAGAAATAGAAAGTTTCAGATCAAGGTTCAGGAAACAAAGAAACAGC 120
Db |||||
61 GGAATAATACATAATCTAGAAATAGAAAGTTTCAGATCAAGGTTCAGGAAACAAAGAAACAGC 120
Qy 121 TGAATACCAACAGGATATCTGTGGTAAGCGGTTCTGCGCCCGCTCAGGCGCCAGGAACA 180
Db |||||
121 TGAATACCAACAGGATATCTGTGGTAAGCGGTTCTGCGCCCGCTCAGGCGCCAGGAACA 180
Qy 181 GATGAGACAGCTGAGTGTATGCGGCCAAACAGAGATATCTGTGGTAAGCGGTTCTGCGCCCGG 240
Db |||||
181 GATGAGACAGCTGAGTGTATGCGGCCAAACAGAGATATCTGTGGTAAGCGGTTCTGCGCCCGG 240
Qy 241 CTCGGGCGCAAGAAACAGATGTCTCCAGATGCGGTCCAGCCCTCAGCAGTTCTTCAGTGAA 300
Db |||||
241 CTCGGGCGCAAGAAACAGATGTCTCCAGATGCGGTCCAGCCCTCAGCAGTTCTTCAGTGAA 300
Qy 301 TCATCAGATGTTTCCAGGGTCCCCAAGGACCTGAAATGACCCCTGTACCTTATTTGAAC 360
Db |||||
301 TCATCAGATGTTTCCAGGGTCCCCAAGGACCTGAAATGACCCCTGTACCTTATTTGAAC 360
Qy 361 TAAACCAATCAGTTCGCTTCTCGCTTCTGTTCCGCGCTTCCGCTCTCCGAGCTCAATAAA 420
Db |||||
361 TAAACCAATCAGTTCGCTTCTCGCTTCTGTTCCGCGCTTCCGCTCTCCGAGCTCAATAAA 420
Qy 421 AGAGCCCAACACCCCTCACTCGGCGCGCCAGTCTTCGATAGACTGCGTCCGCGGGTAC 480
Db |||||
421 AGAGCCCAACACCCCTCACTCGGCGCGCCAGTCTTCGATAGACTGCGTCCGCGGGTAC 480
Qy 481 CCGTATTTCCCAATAAAGCTCTTGGTGTTCGATCCGAATCGTGCTGCTGCTGCTTCTTG 540
Db |||||
481 CCGTATTTCCCAATAAAGCTCTTGGTGTTCGATCCGAATCGTGCTGCTGCTGCTGCTT 540
Qy 541 GGAGGGTCTCTCTGAGTGATTGACTACCAACGAGCGGGGTCTTTTCATTTTGGGGGCTCGT 600
Db |||||
541 GGAGGGTCTCTCTGAGTGATTGACTACCAACGAGCGGGGTCTTTTCATTTTGGGGGCTCGT 600
Qy 601 CCGGGATTTGGAGACCCCTGCGCAGGAGCAACCGACCCACCGCGGAGGTAAAGCTGGCC 660
Db |||||
601 CCGGGATTTGGAGACCCCTGCGCAGGAGCAACCGACCCACCGCGGAGGTAAAGCTGGCC 660
Qy 661 AGCAACTTATCTGTCTGCTCCGATTGCTAGTGTCTATGTTGATGTTATGCGCCTGCG 720
Db |||||
661 AGCAACTTATCTGTCTGCTCCGATTGCTAGTGTCTATGTTGATGTTATGCGCCTGCG 720
Qy 721 TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCGGTGCGAACTGACGAGTT 780
Db |||||
721 TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCGGTGCGAACTGACGAGTT 780
Qy 781 CTGAACACCCCGCGCAACCCCTGGGAGACGTCACAGGACTTTGGGGGCGCTTTTGTGG 840
Db |||||
781 CTGAACACCCCGCGCAACCCCTGGGAGACGTCACAGGACTTTGGGGGCGCTTTTGTGG 840
Qy 841 CCCGACCTGAGGAAGGAGTGCATGTGGAATCCGACCCCGCTCAGGATATGTGGTTCTGGT 900
Db |||||
841 CCCGACCTGAGGAAGGAGTGCATGTGGAATCCGACCCCGCTCAGGATATGTGGTTCTGGT 900
Qy 901 AGGAGACGAGAACCTTAAACAGTTTCCGCGCTCCGCTCTGAAATTTTTCGTTTGGAA 960
Db |||||
901 AGGAGACGAGAACCTTAAACAGTTTCCGCGCTCCGCTCTGAAATTTTTCGTTTGGAA 960
Qy 961 CCGAAGCGCGCGCTTGTGCTGTCGACGCAAGCTTGGGGTGCAGGTCCGACTCTAGAGGA 1020
Db |||||
961 CCGAAGCGCGCGCTTGTGCTGTCGACGCAAGCTTGGGGTGCAGGTCCGACTCTAGAGGA 1020
Qy 1021 TCAATTCGGCACGAGTAAATCGGTGCTGCCGTCTTTAGGACATATGAAGTATGCAAGT 1080
Db |||||
1021 TCAATTCGGCACGAGTAAATCGGTGCTGCCGTCTTTAGGACATATGAAGTATGCAAGT 1080

Qy 1081 GGGATGACTTTCTGTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTGC 1140
Db |||||
1081 GGGATGACTTTCTGTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTGC 1140
Qy 1141 ATGCTCGCTCAGTGACAGCTTTGCTTCTCCCATCTCTAAAAATGCCCCAACTCTTCAAG 1200
Db |||||
1141 ATGCTCGCTCAGTGACAGCTTTGCTTCTCCCATCTCTAAAAATGCCCCAACTCTTCAAG 1200
Qy 1201 AGAGGATGAAGTCTTATAAACTGCACTGATCACTCCCTTTATCTCATTTGTGTTGTAGTTC 1260
Db |||||
1201 AGAGGATGAAGTCTTATAAACTGCACTGATCACTCCCTTTATCTCATTTGTGTTGTAGTTC 1260
Qy 1261 TCGTGCCCATCATTTGGCATAGTGGCAGCTCAGCTCTCTGAAATGCGAAACGAAATTTGCA 1320
Db |||||
1261 TCGTGCCCATCATTTGGCATAGTGGCAGCTCAGCTCTCTGAAATGCGAAACGAAATTTGCA 1320
Qy 1321 CGTTTGGCTCAGTTAATGCGATATATCTCAAAGTCCGGAAGCGCAATGCGCAAGT 1380
Db |||||
1321 CGTTTGGCTCAGTTAATGCGATATATCTCAAAGTCCGGAAGCGCAATGCGCAAGT 1380
Qy 1381 AAGATGAAATGAGATTTGCGAAGAGCTGTGATGGAACGCGATGAGCAACATCGAAAGCAGAA 1440
Db |||||
1381 AAGATGAAATGAGATTTGCGAAGAGCTGTGATGGAACGCGATGAGCAACATCGAAAGCAGAA 1440
Qy 1441 TCCAGTATCTTTCCAGATAATGAAAGCCAACTCTCTAGATGCTAAGAAATTTCCAAAATTTCA 1500
Db |||||
1441 TCCAGTATCTTTCCAGATAATGAAAGCCAACTCTCTAGATGCTAAGAAATTTCCAAAATTTCA 1500
Qy 1501 GCATAACAACCTGATCAAAAGTTAATGATGTTCTTTTCCAGCTTAAATTCCTTACTTTCTCT 1560
Db |||||
1501 GCATAACAACCTGATCAAAAGTTAATGATGTTCTTTTCCAGCTTAAATTCCTTACTTTCTCT 1560
Qy 1561 CCATCCAGGACATGAGNATNATCATAGGGGATATCTCCAGTCTNATAGTAGTCTGGAACA 1620
Db |||||
1561 CCATCCAGGACATGAGNATNATCATAGGGGATATCTCCAGTCTNATAGTAGTCTGGAACA 1620
Qy 1621 CCACAGTACTTTGATTTGCGAGTTTCAAGTATTTGAAACACCTGAAATGCGAGAGTCAAGAGAAATG 1680
Db |||||
1621 CCACAGTACTTTGATTTGCGAGTTTCAAGTATTTGAAACACCTGAAATGCGAGAGTCAAGAGAAATG 1680
Qy 1681 CATTAAAAACAAGAGGAGATCGTAAATTAGAGGAGCGTATATACAATGCAATCAGCAG 1740
Db |||||
1681 CATTAAAAACAAGAGGAGATCGTAAATTAGAGGAGCGTATATACAATGCAATCAGCAG 1740
Qy 1741 AAATTAAGTCTCTAGATGAAAAACAGTATATTTGGAAACAGGAAATAAAAAGGGGAAATGA 1800
Db |||||
1741 AAATTAAGTCTCTAGATGAAAAACAGTATATTTGGAAACAGGAAATAAAAAGGGGAAATGA 1800
Qy 1801 AACTGTTGAATATATACATAATGATCTGAGGCTGAAGGATTTGGGAACTTCTCAGACAT 1860
Db |||||
1801 AACTGTTGAATATATACATAATGATCTGAGGCTGAAGGATTTGGGAACTTCTCAGACAT 1860
Qy 1861 TGAATAATATACATTTTACTCCAAAGGTGCGAGAAAGTGTCTGCTGACTGGGAAATGGACCA 1920
Db |||||
1861 TGAATAATATACATTTTACTCCAAAGGTGCGAGAAAGTGTCTGCTGACTGGGAAATGGACCA 1920
Qy 1921 ACCATCTGGGCTCAACATGACCATCGGGCTGTGAAACAGCAGAGGTGAATTCACAGGCA 1980
Db |||||
1921 ACCATCTGGGCTCAACATGACCATCGGGCTGTGAAACAGCAGAGGTGAATTCACAGGCA 1980
Qy 1981 CCTACATCACAGCGGTAAACAGCCACATCAATGAGATCAAAAGAGTCAACACTGATCGGGA 2040
Db |||||
1981 CCTACATCACAGCGGTAAACAGCCACATCAATGAGATCAAAAGAGTCAACACTGATCGGGA 2040
Qy 2041 CACAAAACACCATCAACAGAGGACCCAGCCACCTTTGGCTTTCACCGTCAATTTGGAAGT 2100
Db |||||
2041 CACAAAACACCATCAACAGAGGACCCAGCCACCTTTGGCTTTCACCGTCAATTTGGAAGT 2100
Qy 2101 TTTTCAAGTCCACACTGTCTTCAAGGCGCAGTCTTATAGACAGGAATGGGAGAGG 2160
Db |||||
2101 TTTTCAAGTCCACACTGTCTTCAAGGCGCAGTCTTATAGACAGGAATGGGAGAGG 2160
Qy 2161 TCCTGAAGACCATGTGGCTGCTCGGCTCAAGTGTAAATGATGATGATGATGATGATGATGATG 2220

Db 2161 TCCGAAAGCATTGTCGCTGCTGCGGTCAAGTGTTAATGCAATTGGTGATGACTGGAAAG 2220
Qy 2221 CTACAGGGTCGGCATCAACATCTTCACTCGCCCTGCGCACAGAGAGAGTGAGTG 2280
Db 2221 CTACAGGGTCGGCATCAACATCTTCACTCGCCCTGCGCACAGAGAGAGTGAGTG 2280
Qy 2281 ACCAAGTCTCTCTGAGTCCAGGTGAAAAGAGAGATAGAGCCCTCTCTGACAAAATGG 2340
Db 2281 ACCAAGTCTCTCTGAGTCCAGGTGAAAAGAGAGATAGAGCCCTCTCTGACAAAATGG 2340
Qy 2341 TATACAGGCTTCCAGGTCTAATAGTCTAAGTCTTAAAGTGATCGGGGGATCT 2400
Db 2341 TATACAGGCTTCCAGGTCTAATAGTCTAAGTCTTAAAGTGATCGGGGGATCT 2400
Qy 2401 CTGGTTTACCTGGAGTTCAGGATTCAGGACCAATGGGGAAGACCGGGAAGCCAGGAC 2460
Db 2401 CTGGTTTACCTGGAGTTCAGGATTCAGGACCAATGGGGAAGACCGGGAAGCCAGGAC 2460
Qy 2461 TTAATGGACAAAAGGCGCAGAGGAGAAAAGGGAGTGGAAGCAATCTTA 2520
Db 2461 TTAATGGACAAAAGGCGCAGAGGAGAAAAGGGAGTGGAAGCAATCTTA 2520
Qy 2521 ATACAGTCCGACTGGTGGGTCGAGCGCCCTCAGAGGCAGAGTGGAGTCTTCAAG 2580
Db 2521 ATACAGTCCGACTGGTGGGTCGAGCGCCCTCAGAGGCAGAGTGGAGTCTTCAAG 2580
Qy 2581 AAGCCAGTGGGTGACGGTGTGACACCGCTCGGAACTCGTGGAGGACTGCTGCT 2640
Db 2581 AAGCCAGTGGGTGACGGTGTGACACCGCTCGGAACTCGTGGAGGACTGCTGCT 2640
Qy 2641 GCAGGACTTTGGGATCAAAAGGTGTTCAAAGTGTGCAATAGCGAGCTTATTTGGAAAAG 2700
Db 2641 GCAGGACTTTGGGATCAAAAGGTGTTCAAAGTGTGCAATAGCGAGCTTATTTGGAAAAG 2700
Qy 2701 GTACGGTCCAAATATGCTGAATGAAGTATTTGTTTCGGGAAAGAGTCATCCATTGAAG 2760
Db 2701 GTACGGTCCAAATATGCTGAATGAAGTATTTGTTTCGGGAAAGAGTCATCCATTGAAG 2760
Qy 2761 AGTCAGAAATPAGACAGTGGGTGTAGAGCTGTTCGACAGCAAGAGATGCTGGGGTC 2820
Db 2761 AGTCAGAAATPAGACAGTGGGTGTAGAGCTGTTCGACAGCAAGAGATGCTGGGGTC 2820
Qy 2821 ACTTTGACCTACATAATGCATATATTTCAATCACAATTTTAAACTGTATTAAGTG 2880
Db 2821 ACTTTGACCTACATAATGCATATATTTCAATCACAATTTTAAACTGTATTAAGTG 2880
Qy 2881 ATTTTTTCTTTGCTTCACTAAATCAGCTTAATTAATTTAAGAACTTAAGAAATTTT 2940
Db 2881 ATTTTTTCTTTGCTTCACTAAATCAGCTTAATTAATTTAAGAACTTAAGAAATTTT 2940
Qy 2941 ATCCACAGAAAGGAATATTTAAATACTCTGGATAAACATATAAATAGCTTCATATTT 3000
Db 2941 ATCCACAGAAAGGAATATTTAAATACTCTGGATAAACATATAAATAGCTTCATATTT 3000
Qy 3001 GCTTCAATAACAGAACCAATTTCACTCTAGGTTTTTAAGTGGCTCGTGGCGGAATG 3060
Db 3001 GCTTCAATAACAGAACCAATTTCACTCTAGGTTTTTAAGTGGCTCGTGGCGGAATG 3060
Qy 3061 ATCCCTCAGGATATAGTATTTGCTTTGCTATAGGAGGGGGAATATGATGCTTATGC 3120
Db 3061 ATCCCTCAGGATATAGTATTTGCTTTGCTATAGGAGGGGGAATATGATGCTTATGC 3120
Qy 3121 AATACTCTGAGTCTGACATGTTAAGTAACTAGTAACTAGTAACTAGTAACTAGTAA 3180
Db 3121 AATACTCTGAGTCTGACATGTTAAGTAACTAGTAACTAGTAACTAGTAACTAGTAA 3180
Qy 3181 GAAAAGCAGCCGTCATGCGGATTTGGTGAAGTAACTAGTAACTAGTAACTAGTAA 3240
Db 3181 GAAAAGCAGCCGTCATGCGGATTTGGTGAAGTAACTAGTAACTAGTAACTAGTAA 3240
Qy 3241 AAGCAACAGACGGGTCTGACATGGAATGGAACCACTGAATTCGCGATTCAGAGAT 3300

Db 3241 AAGCAACAGACGGGTCTGACATGGAATTGGACGAACCACTGAATTCGCGATTCAGAGAT 3300
Qy 3301 ATTGTATTTAAGTGCCTAGCTCGATACAGCAAAACGCCATTTGACATTTACCAATTTGGT 3360
Db 3301 ATTGTATTTAAGTGCCTAGCTCGATACAGCAAAACGCCATTTGACATTTACCAATTTGGT 3360
Qy 3361 GTGCACCTCCAGGTTTCAAGTTCAGCTGCGCAAGCACTCAGGGCGCAAGGGCTGCTAAAGGAAG 3420
Db 3361 GTGCACCTCCAGGTTTCAAGTTCAGCTGCGCAAGCACTCAGGGCGCAAGGGCTGCTAAAGGAAG 3420
Qy 3421 CGGAACACGTAGAAGCAGTCCGACAGAAACGGTGTGACCCCGATGAATGTCAAGCTAC 3480
Db 3421 CGGAACACGTAGAAGCAGTCCGACAGAAACGGTGTGACCCCGATGAATGTCAAGCTAC 3480
Qy 3481 TGGGCTATCTGGAACAAGGAAAACGCAAGCGCAAGAGAAAGCAAGTAGCTTGCAGTGGG 3540
Db 3481 TGGGCTATCTGGAACAAGGAAAACGCAAGCGCAAGAGAAAGCAAGTAGCTTGCAGTGGG 3540
Qy 3541 CTTCATGCGGATAGCTAGACTGGGCGGTTTATATGGAACAGCAAGCAACCGGAATTTGCCA 3600
Db 3541 CTTCATGCGGATAGCTAGACTGGGCGGTTTATATGGAACAGCAAGCAACCGGAATTTGCCA 3600
Qy 3601 GCTGGGCGGCTCTGCTGAAGTTTGGGAAGCCCTGCAAAAGTAAACTGATGGCTTTCTTG 3660
Db 3601 GCTGGGCGGCTCTGCTGAAGTTTGGGAAGCCCTGCAAAAGTAAACTGATGGCTTTCTTG 3660
Qy 3661 CCGCAAGAGTCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGGATGAGATCG 3720
Db 3661 CCGCAAGAGTCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGGATGAGATCG 3720
Qy 3721 TTTGCGATGATGAAAGATGGAATGCAACGAGTTCTCGGCGCTTTCGGTGGAGAGG 3780
Db 3721 TTTGCGATGATGAAAGATGGAATGCAACGAGTTCTCGGCGCTTTCGGTGGAGAGG 3780
Qy 3781 CTATTCGGCTATGACTGGGCACACACAACTGGCTGCTCTGATGCGCGCTGTTCCGG 3840
Db 3781 CTATTCGGCTATGACTGGGCACACACAACTGGCTGCTCTGATGCGCGCTGTTCCGG 3840
Qy 3841 CTGTACGCGAGGGGCGCCGCTTCTTTTGTCAAGACCGACCTGTGCGGTGCCCTGAAT 3900
Db 3841 CTGTACGCGAGGGGCGCCGCTTCTTTTGTCAAGACCGACCTGTGCGGTGCCCTGAAT 3900
Qy 3901 GAATGCGAGACGAGGACGCGGCTATCGTGGCTGGCCACGACGCGGCTTCTTTGGCA 3960
Db 3901 GAATGCGAGACGAGGACGCGGCTATCGTGGCTGGCCACGACGCGGCTTCTTTGGCA 3960
Qy 3961 GCTGTCTCGACCTTGTCACTGAAGCGGAAGGAGCTGGCTGCTATTTGGGCGAAGTCCG 4020
Db 3961 GCTGTCTCGACCTTGTCACTGAAGCGGAAGGAGCTGGCTGCTATTTGGGCGAAGTCCG 4020
Qy 4021 GGGCAGGATCTCTGTCTCATCTCACCTTGTCTCTGCGCGAGAAAGTATCCATCATGGCTGAT 4080
Db 4021 GGGCAGGATCTCTGTCTCATCTCACCTTGTCTCTGCGCGAGAAAGTATCCATCATGGCTGAT 4080
Qy 4081 GCAATGCGGCTGTCATACGCTTGTATCCGCTACCTGCGCATTCGACCAACAGCGAAA 4140
Db 4081 GCAATGCGGCTGTCATACGCTTGTATCCGCTACCTGCGCATTCGACCAACAGCGAAA 4140
Qy 4141 CATCGCATCGAGCAGCAGCTACTCGGATGGAAGCCGCTTGTGCTGATCAGGATGATCTG 4200
Db 4141 CATCGCATCGAGCAGCAGCTACTCGGATGGAAGCCGCTTGTGCTGATCAGGATGATCTG 4200
Qy 4201 GACGAGAGCATCAGGGCTCGGCGCAGCGCAACTGTTTGGCAGGCTCAAGCGCGCATG 4260
Db 4201 GACGAGAGCATCAGGGCTCGGCGCAGCGCAACTGTTTGGCAGGCTCAAGCGCGCATG 4260
Qy 4261 CCGCAGCGGAGGATCTGTCGTAACCCATGGGATGCTGCTTTGCCGAATATCATGGTG 4320
Db 4261 CCGCAGCGGAGGATCTGTCGTCGTAACCCATGGGATGCTGCTTTGCCGAATATCATGGTG 4320
Qy 4321 GAAAATGGCGCTTTTCTGGAATTCATGATGCTGGCGGCTGGGTGGCGGACCGCTAT 4380
Db 4321 GAAAATGGCGCTTTTCTGGAATTCATGATGCTGGCGGCTGGGTGGCGGACCGCTAT 4380

Qy 4381 CAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCAATGGCTGAC 4440
Db |||||
Qy 4381 CAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCAATGGCTGAC 4440
Db |||||
Qy 4441 CGCTTCTCGTCTTTAGCGTATCGCGCTCCCGATTCGCGAGCGATCGCTTCTATCGC 4500
Db |||||
Qy 4441 CGCTTCTCGTCTTTAGCGTATCGCGCTCCCGATTCGCGAGCGATCGCTTCTATCGC 4500
Db |||||
Qy 4501 CTCTTGACGAGTCTTCTGAGCGGACTCGGCGTTCGATAAAATAAAGATTTATTT 4560
Db |||||
Qy 4501 CTCTTGACGAGTCTTCTGAGCGGACTCGGCGTTCGATAAAATAAAGATTTATTT 4560
Db |||||
Qy 4561 AGTCTCCAGAAAAAGGGGGGAATGAAGACCCCACTGTAGTGTGGCAAGCTAGCTTAA 4620
Db |||||
Qy 4561 AGTCTCCAGAAAAAGGGGGGAATGAAGACCCCACTGTAGTGTGGCAAGCTAGCTTAA 4620
Db |||||
Qy 4621 GTAACGCCATTTTGAAGCGATGGAATAATACATACTGAGAAATAGAGAGTTTCAGATCA 4680
Db |||||
Qy 4621 GTAACGCCATTTTGAAGCGATGGAATAATACATACTGAGAAATAGAGAGTTTCAGATCA 4680
Db |||||
Qy 4681 AGGTCAGGAACAGATGGAACAGCTGAAATATGGGCCCAACAGGATATCTGTGTAAGCAGT 4740
Db |||||
Qy 4681 AGGTCAGGAACAGATGGAACAGCTGAAATATGGGCCCAACAGGATATCTGTGTAAGCAGT 4740
Db |||||
Qy 4741 TCCTGCCCGGCTCAGGGCCCAAGAACAGATGGAACAGCTGAAATATGGGCCCAACAGGAT 4800
Db |||||
Qy 4741 TCCTGCCCGGCTCAGGGCCCAAGAACAGATGGAACAGCTGAAATATGGGCCCAACAGGAT 4800
Db |||||
Qy 4801 TCTGTGTGAAGCAGTTCTGCGCGGCTCAGGGCCCAAGAACAGATGTTCCCGAGATCGG 4860
Db |||||
Qy 4801 TCTGTGTGAAGCAGTTCTGCGCGGCTCAGGGCCCAAGAACAGATGTTCCCGAGATCGG 4860
Db |||||
Qy 4861 TCAGGCTCAGCAGTTCTTAGAGAACCATCAGATGTTTCCAGGGTCCCGCCAGGACCTG 4920
Db |||||
Qy 4861 TCAGGCTCAGCAGTTCTTAGAGAACCATCAGATGTTTCCAGGGTCCCGCCAGGACCTG 4920
Db |||||
Qy 4921 AAATGACCTGTGCTTATTTGAACCAACCAATCAGTTCGGTCTCTCGCTTCTGTCGCGC 4980
Db |||||
Qy 4921 AAATGACCTGTGCTTATTTGAACCAACCAATCAGTTCGGTCTCTCGCTTCTGTCGCGC 4980
Db |||||
Qy 4981 GCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACACCCCTCACTCGGGGGCCAGTCCT 5040
Db |||||
Qy 4981 GCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACACCCCTCACTCGGGGGCCAGTCCT 5040
Db |||||
Qy 5041 CCGATTGACTGAGTCGCGCGGTACCGGTGATCCCAATTAACCCCTCTTGCAGTTGCAATCC 5100
Db |||||
Qy 5041 CCGATTGACTGAGTCGCGCGGTACCGGTGATCCCAATTAACCCCTCTTGCAGTTGCAATCC 5100
Db |||||
Qy 5101 GACTTGTGGTCTCGCTGTTCCTTGGAGGGTCTCTCTGAGTGATGACTACCCGTGAGC 5160
Db |||||
Qy 5101 GACTTGTGGTCTCGCTGTTCCTTGGAGGGTCTCTCTGAGTGATGACTACCCGTGAGC 5160
Db |||||
Qy 5161 GGGGGTCTTTCAATTTGG 5177
Db |||||
Qy 5161 GGGGGTCTTTCAATTTGG 5177
Db |||||

RESULT 2
AX018983
LOCUS AX018983 5177 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 1 from Patent WO9942577.
ACCESSION AX018983
VERSION AX018983.1 GI:10043078
KEYWORDS
SOURCE
ORGANISM
synthetic construct
other sequences; artificial sequences.
REFERENCE
1
AUTHORS Yla-Herttuala, S., Airenne, K., Kulomaa, M., Marjomaki, V. and
Lehtolainen, P.
TITLE Biotin-binding receptor molecules
JOURNAL Patent: WO 9942577-A 1 26-AUG-1999;

YLA HERTTUALA SEPPO (FI); AIRENNE KARI (FI); KULOMAA MARKKU (FI);
MARJOMAKI VARP (FI); LEHTOLAINEN PAULIINA (FI); EUROGENE LIMITED
(GB)
FEATURES
source
Location/Qualifiers
1. 5177
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Designed DNA sequence to encode transmembrane
protein."
1071..2273
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC07722.1"
/db_xref="GI:10043079"
/translation="MAQMDFFPQQEDTSDCTESVKFDRASVTALLPPHKNPTLQRE
RMKSYKTAITLILYIVFVLIPIIGIVAAQLKWKNTCTVGSVNADISPSPEKGNG
SEDEMPREAVMERMSNMESRIQYLSNDEANLLDAKNQFNFSITDFNDVLPQNS
LLESIQHENIIGDISLVGLNTVLDLQFSIETLNGRVQENAFKQOENRKLBERI
YNASAEIKSIDKQVYLEQRIKGMKLNNITNDLRLLKDWHSOTLKNITLLQGARIC
SLTGKNTDLSNMTIGAVNSRGEFTGTIYATVATSNIEIKESPLHGTQNTINKTQP
TLGFTVNWKFSESTTFTGQCFIDRNGKEVLKTMWLLURSSVNDIGDDWKATRVGINIF
TLRLTQKE"

ORIGIN

Query Match 100.0%; Score 5177; DB 6; Length 5177;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTTGAAGACCCACCCGTAAGTGGCAAGCTAGCTTAAGTAACGCCCACTTTTGAAGGCAT 60
Db |||||
Qy 1 TTTGAAGACCCACCCGTAAGTGGCAAGCTAGCTTAAGTAACGCCCACTTTTGAAGGCAT 60
Db |||||
Qy 61 GGAATAATACATTAAGTGAATAGAAAAGTTTCAGATCAAGTCAAGAACAAAGAACAGC 120
Db |||||
Qy 61 GGAATAATACATTAAGTGAATAGAAAAGTTTCAGATCAAGTCAAGAACAAAGAACAGC 120
Db |||||
Qy 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTCCTGCCCGCTCAGGGCCCAAGAAC 180
Db |||||
Qy 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTCCTGCCCGCTCAGGGCCCAAGAAC 180
Db |||||
Qy 181 GATGAGACAGCTGAGTGATGGGCCAAACAGGATATCTGTGTAAGCAGTTCCTGCCCGG 240
Db |||||
Qy 181 GATGAGACAGCTGAGTGATGGGCCAAACAGGATATCTGTGTAAGCAGTTCCTGCCCGG 240
Db |||||
Qy 241 CTGGGGCCAAAGAACAGATGGTCCCGAGTGGCGGTCCAGCCCTCAGAGTTCTAGTGAA 300
Db |||||
Qy 241 CTGGGGCCAAAGAACAGATGGTCCCGAGTGGCGGTCCAGCCCTCAGAGTTCTAGTGAA 300
Db |||||
Qy 301 TCATCAGATGTTTCCAGGGTGGCCCAAGGACCTGAAATGACCTTATTTGAAC 360
Db |||||
Qy 301 TCATCAGATGTTTCCAGGGTGGCCCAAGGACCTGAAATGACCTTATTTGAAC 360
Db |||||
Qy 361 TAAACCAATCAGTTCGCTTCTGCTTCTGTTGGCGGTTCCTCGGTCTCCGAGCTCAATAA 420
Db |||||
Qy 361 TAAACCAATCAGTTCGCTTCTGCTTCTGTTGGCGGTTCCTCGGTCTCCGAGCTCAATAA 420
Db |||||
Qy 421 AGAGCCCAACCCCTCACTCGGGCGGCCAGTCTTCGATAGACTGCGTCCCGGGGTAC 480
Db |||||
Qy 421 AGAGCCCAACCCCTCACTCGGGCGGCCAGTCTTCGATAGACTGCGTCCCGGGGTAC 480
Db |||||
Qy 481 CCGTATTTCCCAATAAAGCCTCTTGTGTTTGCATCCGAATCGTGGTCTCGCTGTTCTT 540
Db |||||
Qy 481 CCGTATTTCCCAATAAAGCCTCTTGTGTTTGCATCCGAATCGTGGTCTCGCTGTTCTT 540
Db |||||
Qy 541 GAGGGGTCTCTCTGAGTGATTCAGTACCCCAAGCGGGGTCTTTTCAATTTGGGGGCTCGT 600
Db |||||
Qy 541 GAGGGGTCTCTCTGAGTGATTCAGTACCCCAAGCGGGGTCTTTTCAATTTGGGGGCTCGT 600
Db |||||
Qy 601 CCGGGATTTGGAGACCCCTGCCAGGGACCAACCGACCAACCGGGAGGTAACTGGCC 660
Db |||||
Qy 601 CCGGGATTTGGAGACCCCTGCCAGGGACCAACCGACCAACCGGGAGGTAACTGGCC 660
Db |||||

Qy 5041 CCGATTGACTGAGTCGCCCGGGTACCCTGATATCCAAATAAACCTCTTTGCAAGTTGCATCC 5100
Db 5041 CCGATTGACTGAGTCGCCCGGGTACCCTGATATCCAAATAAACCTCTTTGCAAGTTGCATCC 5100
Qy 5101 GACTTGTGGTCTCGCTGTTCTTGGAGGGTCTCTCTGAGTGATGACTACCGGTGAGC 5160
Db 5101 GACTTGTGGTCTCGCTGTTCTTGGAGGGTCTCTCTGAGTGATGACTACCGGTGAGC 5160
Qy 5161 GGGGGTCTTTTCATTGG 5177
Db 5161 GGGGGTCTTTTCATTGG 5177

RESULT 3

LOCUS A80474 7699 bp DNA circular PAT 21-JAN-2000
DEFINITION Sequence 1 from Patent WO9943795.
ACCESSION A80474
VERSION A80474.1 GI:6731328
KEYWORDS
SOURCE
ORGANISM unidentified
unclassified sequences.
REFERENCE 1 (bases 1 to 7699)
AUTHORS Benkel,B.F. and Falconer,M.M.
TITLE TRANS-SOMATICS WITH GENE TRANSFER INTO MAMMARY EPITHELIAL CELLS
JOURNAL Patent: WO 9943795-A 1 02-SEP-1999;
BENKEL BERNHARD F (CA); FALCONER MARCIA M (CA)
FEATURES
source
1..7699
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 40.2%; Score 2078.8; DB 6; Length 7699;
Best Local Similarity 64.3%; Pred. No. 0;
Matches 3339; Conservative 0; Mismatches 1814; Indels 38; Gaps 14;
Qy 1 TTTGAAAGACCCACCGTAGGTGGCAGCTAGCTTAAGTAACGCCACTTTGCAAGGCAT 60
Db 175 TTTGAAAGACCCACCGTAGGTGGCAGCTAGCTTAAGTAACGCCACTTTGCAAGGCAT 234
Qy 61 GGAATAATACATACTGAGATAGAAAAGTTTCAGATCAAGGTCAAGGACAAAGAAACAGC 120
Db 235 GGAATAATACATACTGAGATAGAAAAGTTTCAGATCAAGGTCAAGGACAAAGAAACAGC 294
Qy 121 TGAATACAAACAGAGATATCTGTGTAAGCGGTTCCTGCCCCGCTCAGGCCCAAGAAC 180
Db 295 TGAATACAAACAGAGATATCTGTGTAAGCGGTTCCTGCCCCGCTCAGGCCCAAGAAC 354
Qy 181 GATGAGACAGCTGATGATGGCCCAACAGAGATATCTGTGTAAGCGGTTCCTGCCCCG 240
Db 355 GATGAGACAGCTGATGATGGCCCAACAGAGATATCTGTGTAAGCGGTTCCTGCCCCG 414
Qy 241 CTCGGGGCAAGAACAGATGTTCCCAAGTCCAGTCCAGCCCTCAGCAGTTTCTAGTGAA 300
Db 415 CTCGGGGCAAGAACAGATGTTCCCAAGTCCAGTCCAGCCCTCAGCAGTTTCTAGTGAA 474
Qy 301 TCATCAGATGTTTCCAGGGTCCCAAGGACCTGAAAATGACCTGTACCTTATTTGAAC 360
Db 475 TCATCAGATGTTTCCAGGGTCCCAAGGACCTGAAAATGACCTGTACCTTATTTGAAC 534
Qy 361 TAACCAATCAGTTGCTTCTGCTTCTGTTTCCGGGGCTTCGGCTCTCGAGCTCAATAAA 420
Db 535 TAACCAATCAGTTGCTTCTGCTTCTGTTTCCGGGGCTTCGGCTCTCGAGCTCAATAAA 594
Qy 421 AGAGCCCAACACCTCTACTCGGGCGCCAGTCTTCGATAGACTGCGTCCCGGGTAC 480
Db 595 AGAGCCCAACACCTCTACTCGGGCGCCAGTCTTCGATAGACTGCGTCCCGGGTAC 654
Qy 481 CCGTATTTCCCAATAAAGCCTCTTGTGTTTGCATCCGAATCGTGTCTCGCTGTTCTTGG 540

Db 655 CCGTATTTCCCAATAAAGCCTCTTGTGTTTGCATCCGAATCGTGTCTCGCTGTTCTTGG 714
Qy 541 GAGGGGTCTCTCTCTGAGTGATTTGACTACCAACGACGGGGGTCTTTTCATTTTGGGGGCTCGT 600
Db 715 GAGGGGTCTCTCTCTGAGTGATTTGACTACCAACGACGGGGGTCTTTTCATTTTGGGGGCTCGT 774
Qy 601 CCGGGATTTGGAGACCCCTGCGGACGACCAACGACCCACCGGAGGTAACTGGCC 660
Db 775 CCGGGATTTGGAGACCCCTGCGGAGGACCAACGACCCACCGGAGGTAACTGGCC 834
Qy 661 AGCAACTTATCTGTGTCGTGTCGATTTCTAGTGTCTATGTTTATGTCCTGCGCTGCG 720
Db 835 AGCAACTTATCTGTGTCGTGTCGATTTCTAGTGTCTATGTTTATGTCCTGCGCTGCG 894
Qy 721 TCTGTACTAGTTAGTAACTAGCTCTGTATCTGCGGACCCGCTGGTGTGAACTGACGAGTT 780
Db 895 TCTGTACTAGTTAGTAACTAGCTCTGTATCTGCGGACCCGCTGGTGTGAACTGACGAGTT 954
Qy 781 CTGAACACCCGGCCGCAACCCCTGGGAGACGTCCAGGACCTTTGGGGGCCCTTTTGTGG 840
Db 955 CTGAACACCCGGCCGCAACCCCTGGGAGACGTCCAGGACCTTTGGGGGCCCTTTTGTGG 1014
Qy 841 CCGGACCTGAGGAGGAGTCTGATGTGGAATCCGACCCCGTCAGATATGTTCTGCT 900
Db 1015 CCGGACCTGAGGAGGAGTCTGATGTGGAATCCGACCCCGTCAGATATGTTCTGCT 1074
Qy 901 AGGAGACGAGAACCTAAACAGTTCCCGCTCTCTGAAATTTTGTCTTTCGGTTTGGAA 960
Db 1075 AGGAGACGAGAACCTAAACAGTTCCCGCTCTCTGAAATTTTGTCTTTCGGTTTGGAA 1134
Qy 961 CCGAAGCCGGCGCTTTGTCTGTGTCAGCCAGCTTTGGGTCGAGGTGCACTCTAGAGGA 1020
Db 1135 CCGAAGCCGGCGCTTTGTCTGTGTCAGCGCTGCAGCATCGTTCTGTGTTCTCTGTCT 1194
Qy 1021 TCAATTCGGACGAGTAAATCGGTGCTGCCGTCTTTAGGACATATGAAATGSCACAGT 1080
Db 1195 GACTGTGTTTCTGTGTTTGTCTGAAATTTAGGGCCAGACTGTTTACCCTCCCTTAAGTTT 1254
Qy 1081 GGGATGACTTTTCTGTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTCG 1140
Db 1255 GACCTTAGGTACATGGAAGATGTCGAGCGGATCGCTCACACCGCTGGTAGATGCA 1314
Qy 1141 ATGCTCGCTCAGTGACAGCTTTGCTTCTCCCCATCTTAAATAATGAGGCGCACTCTTCAAG 1200
Db 1315 GAAGAGAGCTTTGGGTACCTTCTGCTCTGCAAGATGCGCAACCTTTAACTGCGATGGCC 1374
Qy 1201 AGAGGATGAGTCTTATTAACCTGCACTGATCACCCTTTATCTCATTTGTGTTGAGTTC 1260
Db 1375 GCGAGACGGCACCTT-TAACCGAGACCTCATCACCCAGGTTAAGATCAAGGTCTT--TTC 1431
Qy 1261 TCGTGCCCATCATTTGGCATAGTGGCAGCTCAGCTCTCTGAAATGGGAAACGAAGAATTGCA 1320
Db 1432 ACCTGGCCGATGAGACACCCAGACCGAGTCCCTCATCTGCTGACCTGGGAGGCTTGGC 1491
Qy 1321 CGGTGGCTCAGTTAATGACAGATATCTCCAAGTCCGGAAGGCAAGAAATGCGAGTG 1380
Db 1492 TTTTGACCCCTCTCCCTGGGTCAAGCCCT----TTGTACACCTTAAGCTCCGCTCCT 1546
Qy 1381 AAGATGAATGAGATTTGAGAGAGCTGTGTAAGCGATGGAACCAATGGAAGACGAA 1440
Db 1547 CTTCTCCATCCGCCCCGCTCTCTCCCTTTGAACTCTCGTTGACCCCGCTCGATCC 1606
Qy 1441 TCCAGTATTTTCAGATAATGAAGCCAAATCTCTPAGATGCTAAGAATTTCCAAAATTTCA 1500
Db 1607 TCCCTTTATCAGGCCCTCATCTCTCTTCTAGGCGCCCGGAATTCGTTAACTCGACATGAA 1666
Qy 1501 GCATAACAACTGATCAAGATTTAATGATGTTCTTTTCCAGCTAAATTTCTTACTTTCTCT 1560
Db 1667 GTCTTCTCTCTCGCAGCTGTGGGCTTTGCTGGGCAAGTACAAATCCCAACACTCAG 1726
Qy 1561 CCATCCAGG--AACATGAGAAATATCATAGGGGATATCTCCAAGTCAATAGTAGTCTGAA 1618

Db 1727 GCTGGAGGACATCTATCGTGCACTCTCTTTGAATGGCGCTGGCGGACATTCGCACTGGAG 1786
Qy 1619 CACCACAGTACTTGATTTTGCAGTTTCAGTATTTGAAACACTGAATGCGAGATCCCAAGAGAA 1678
Db 1787 TGGGAACACTATTTAGCTCTTAATGGGTTTGGAGAGTTTCAGGTTTCTCTCCAAATGAA 1846
Qy 1679 TGCATTTTAAACAAACAAAGAGAGATGCGTAAATTTAGAGGAGCGTATATACATGATCATCAGC 1738
Db 1847 AACATGTCTATTAATCCGAAACAGGCCCTGGTGGGAAAGATACCAGCCCATCAGCTAC 1906
Qy 1739 AGAAATTAAGTCTCTAGATGAAACAAAGTATATTTGGAAACAGGAAATAAAGGGGAAAT 1798
Db 1907 AAGATCTGCAGTCGATCGGCAATGAAATGAAATTTTCAAGAGACATGGTGACAGATGCAAC 1966
Qy 1799 GAAACTGTTGAAATATATACATTAATGATCTGAGGCTGAGGATTTGGGNAACATTTCTCAGAC 1858
Db 1967 AATGTTGGAG--TTGATTTTATGTGGATGCTGTGTGTCATCACAATGTGTGGATCTATGGG 2025
Qy 1859 ATTGAAATATATCACTTTACTTCAAGGTGCCAGAAAGTCTCGCTGACTGGGAAATGGAC 1918
Db 2026 TGGCAGGGCACCCACTCAACATGTGGAGCTATTTCAACACCGGGACTAGAGATTTTCC 2085
Qy 1919 CAACGATCTGGGCTCCAAACATGACATCGGGGCTGTGAAACAGCAGAGGTGAATTCACAGG 1978
Db 2086 CGCTGTGCGTACTCTGCTGGGATTTCAATGACGGCAAAATGTCACTGCAAGTGGAGA 2145
Qy 1979 CACCTACATCACAGCGTTAACAGCCACATCAATGAGATCAAGAGTCAACACATTCGATGG 2038
Db 2146 CATCGAAATTTATGGGGACATGTATCAGGTCGGGATTTGCAAGTTGTCCAGCCTTCTTGA 2205
Qy 2039 GACA--CAAAAACCAATCAACAAAGAGGCCACAGCCACCTTTGGCTTCACCGTCAATTTGA 2097
Db 2206 TCTGGCTCTGGAGAGGACTATGTACGCTCAACAAATTCGACGCTACATGAATCACCTCAT 2265
Qy 2098 AGTTTTCAGATCCACCATCTCTTTCACGGGCCAGTGTCTCATAGACAGGAATGGGAAGG 2157
Db 2266 TGATATGGGTAGCAGGGTTCCGGATCGATGCTGCCAAGCATGTGGCCAGGGGACAT 2325
Qy 2158 AGGTCCTGAAGACCATGTGGCTGTGGCTCAAGTGTATGATGATTTGGTGTGATCTGGA 2217
Db 2326 AAGAGCGTTTCTGACAAACTGACAGATCTAAATACTCAGTGGTTTTGACAGGAACGAA 2385
Qy 2218 AAGCTACAGGGTC--GGCATCAACATCTTCACTCGCCTCGCCACACAGAGAGAGTGAGTG 2276
Db 2386 ACCCTTTATTTACCAAGAGTAAATGTACTTGGGAGGAGAGCCATCACAGGCACTCAGTA 2445
Qy 2277 AGTACCAAGTCTCTCTG---GACTCCAGGTGAAAGAGAGATAGAGGCCCTCTCGGAC 2333
Db 2446 CTTTGGGAAATGGCCGCGTGACAGAAATTCAGTATGGTGTGCCAATCTGGGGACGGTGATCCG 2505
Qy 2334 AAAATGGTATACAGGCTTTCCAGGCTTAATAGTACTTCCAGGCTTTAAAGGTGATCGGG 2393
Db 2506 GAAGTGAATGGAGAGATGGGCTACTTAAAGAACTGGGGAGAAAGGCTGGGGCTTTGT 2565
Qy 2394 GGGATCTCTGTTTACCTTGGAGTTTCGAGATTTCCAGGAC--CAATGGGGAAGACCCGGA 2451
Db 2566 GCCTTCTGACAGAGCCCTGTGTTTGTGGATTAACAACAGAACCCAGCGGGGGCAGGGGC 2625
Qy 2452 AGCCAGGACTTAAATGGACAAAAGGCCAGAGAGGAGAAAGAGGAGTGGAGCATGCAAA 2511
Db 2626 AGCGGAGCTTCCATCTTTACTTTCTGGATGCGCAGGCTTTTATAAATGGCGGTGGTTTT 2685
Qy 2512 GACAACTAATACAGTCCGACTGGTGGGTGGGAGCGGCCCTTCACGAAGGCAGAGTGAGA 2571
Db 2686 CATGCTCGCTCATCCGTACCGGTTTCAACCGGTTGATGTCAAGTTTATCGTTGGCCAAAGATA 2745
Qy 2572 TTTTTCAGAGGCGCAGTGGGTTACGGTGTGTGACGACCGCTGGGAACTGGGTGGAGGAC 2631
Db 2746 TTTTCGAAACCGAGTGGATGTATACGACTGGGTGGGACCAACAAAGTAACTCGGACGATC 2805
Qy 2632 T----GFTGCTGACAGGACTTGGGATPACAAAGGTGTTCAAAGGTGTTCAAAGTGTGCAATAAGCGAGCT 2687
Db 2806 GACGAAGTCCGTTTAAATCAACGACGACACATACCTGTGGCAATGACTGGGTCTGCGAACA 2865

Qy 2688 TATTTTGGAAAAAGTACGGGTCCAATATGGCTGAATGAAGTATTTTGTTCGGGAAAGAG 2747
Db 2866 TCGTGGCGCAAAATAAG---AACATGGTTATCTTCCGTAATGTGGTAGACCGTCAAG 2921
Qy 2748 TCATCCATTTGAAGAGTGCAGAAATTTAGACGTGGGGTGTGAGAGCCTGTTCGACAGAGAA 2807
Db 2922 CTTTCTCAAACTGGTGGGACAAACGGGAGCAATCAAGTAGCTTTTGGTTCGGCGGACAGAG 2981
Qy 2808 GATGCTGGGGTCACTTTGTCACCTACATAATGCAATCATATTTTTCATTTACATTTTTTAA 2867
Db 2982 GCTTCATTTGTCTTTAATAATGATGACTGGTATATGAATTCGATTTTGCAACTGGTCTGC 3041
Qy 2868 CTGTTATAAAGTGTATTTTTTCTTTCCTTTCCTTCACTTAAATCAGCTTAATTAATAATTAAGA 2927
Db 3042 CTGCTGNAACCTACTCGGATGTTTATTTCTGGACAAAGGAGGAGCGTGGTGTACTGGAA 3101
Qy 2928 AACTAAGAAATTTTATCCACAGAAAAGGAATATTTTAAAAATCACTGGATAAACATATAAAA 2987
Db 3102 AGCAGGTGTACGTTTCTCGGATGGAAAGGCCAATTTCCAGATTTAGTAAACAGCGATGAAG 3161
Qy 2988 TAGTTTCATATTTGCTTCAAAATACCAGAACCATTTTCACTTCTCTAGGTTTTTAAGTGGC 3047
Db 3162 ATCCATTTGTGCAATTTACGTTTGAATGCAAGTTTATAAGCTTCGAGGATCCACTAGTAAC 3221
Qy 3048 TCGTGGCGAAATGATCCCTCAGGATATAGTGTTCGCTTTTGTGCAATAGGAGGGGAAA 3107
Db 3222 GCGCGCAGTGTCTGGAATTCG-----CTTGTGCACTCTAGGGCGGCAATTCGCC 3276
Qy 3108 TGTAGTCTTATGCAATACTCTTGTAGTCTTTCGCAACATGTTAAACGATGATTTAGCAATG 3167
Db 3277 CCTCTCCCCCCCCCTTAACGTTTACTGGCGAAGCGCTTGGAAATAAGGCCGTGTGTG 3336
Qy 3168 CTTTACAGAGAGAAAACACCGTGCATGCCGATTTGGTGGAGTAAAGTGGTACGATC 3227
Db 3337 TTTGTCTATATGTGATTTTCCACCATATTCGCGCTTTTGGCAA-----TGTGAGGGCC 3390
Qy 3228 GTGCTTTATTAGGAAGCAACAGACGGGTCTGACATGGATTTGGACGAACCACTCAATTC 3287
Db 3391 CGAAACCTGGCCGTGTCTTTTGACGAGCATTTCTAGGGGTCTTTCCCTCTCGCCAA 3450
Qy 3288 GCATTTGAGAGATATTTGATTTTAAAGTGCCTAGTCTGATACAGCAAAACCCATTTGACCAT 3347
Db 3451 GGAATGCAAGTCTGTTGTAATGCTGGAAGAGCAGTCTCTCTGGAAGCTTCTTTGAAGA 3510
Qy 3348 TCACCAATTTGGTGTGCACTCCAGCTTCACGTCGCCGACGACATCAGCGGCCAAGGG 3407
Db 3511 CAACCAACGCTCTGTAGCGACCTTTTGACGACGCGGAACCCCCACCTGGCGACAGGTGC 3570
Qy 3408 CTGCTAAAGGAAGCGGAACACGCTAGAAAGCGAGTCCGACAGAAACGGTGTGACCCCGAT 3467
Db 3571 CTCTGGGCCAAAGCCACGCTGATATAGATACACTGCAAGGGCGGCAACCCAGTGC 3630
Qy 3468 GAATGTCACTACTGGGCTATCTGGACA--AGGGAACCGCAAGCGCAAGAGAAAGCAGG 3526
Db 3631 CACGTTGTGAGTTGGATAGTTGTGGAAGAGTCAATGGCTCTCTCAAGGTAAGTCAAC 3690
Qy 3527 TAGTTGACGTGGGCTTACATGGCGATAGCTAGACTGGGCGGTTTTTATGACAGCAAGCG 3586
Db 3691 AAGGGGTGAAGGATGCCAAGAGGTACCCCATTTGTATGGGAATCTGATCTGGGGCCCTCG 3750
Qy 3587 AACCGGAATTTGCCAGCTGGGGCGCCCTCTGCTGAAGTTGGGAAGCCCTGCAAAAGTAAACT 3646
Db 3751 GTGCATATGCTTTACATGTGTTTAGTCAGAGTTTAAAAAGCTCTAGGGCCCCCGAACAC 3810
Qy 3647 GGAATGGCTTTTTCGCCCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGA 3706
Db 3811 GGGGACGTGCTTTTCTTTTGAAGAACACGATGATAGCTTTGCCCAACACCCGGGATAATTC 3870
Qy 3707 CAGGATGAGGATCGTTTCGATGATTTGAAACAGATGGAATGACGCGAGGTTCTCCGGCCG 3766
Db 3871 CTGACGCAATATATGGGATCGGCCAATTGAACAGATGGAATGACGCGAGGTTCTCCGGCCG 3930

3767 CTTGGTGGAGAGGCTATTTCGGCTATGACTGGGCACAAACAGACAATCGGCTGCTCTGATG 3826
3931 CTTGGTGGAGAGGCTATTTCGGCTATGACTGGGCACAAACAGACAATCGGCTGCTCTGATG 3990
3827 CCGCGGTGTTCCGGCTGTGACGCGAGGGGCCCGGTTCTTTTGTCAAGACCGACCTGT 3886
3991 CCGCGGTGTTCCGGCTGTGACGCGAGGGGCCCGGTTCTTTTGTCAAGACCGACCTGT 4050
3987 CCGGTGCGCTGAATGAATGACGAGACGAGGCGAGCGGGCTATCTGGGTGGCCACGACGG 3946
4051 CCGGTGCGCTGAATGAATGACGAGACGAGGCGAGCGGGCTATCTGGGTGGCCACGACGG 4110
3947 GCGTTCCTTGGCGAGCTGTCTCGACGTTGTCACGTAAGCGGGAAGGACTGGCTGCTAT 4006
4111 GCGTTCCTTGGCGAGCTGTCTCGACGTTGTCACGTAAGCGGGAAGGACTGGCTGCTAT 4170
4007 TGGCGAAGTCCCGGGCGAGGATCTCTGTCTCATCTCACCTTGTCTCTCGCGGAGAAAGTAT 4066
4171 TGGCGAAGTCCCGGGCGAGGATCTCTGTCTCATCTCACCTTGTCTCTCGCGGAGAAAGTAT 4230
4067 CCATCATGGCTGATGCAATGCGGGCGGTGCATACGCTTGATCGGCTACCTGCGCATTCG 4126
4231 CCATCATGGCTGATGCAATGCGGGCGGTGCATACGCTTGATCGGCTACCTGCGCATTCG 4290
4127 ACCACCAAGCAACATCGATCGAGCGAGCACGCTACTCGGATGGAAGCCGCTTCTGCG 4186
4291 ACCACCAAGCAACATCGATCGAGCGAGCACGCTACTCGGATGGAAGCCGCTTCTGCG 4350
4187 ATCAGGATGATCTGCAAGAGAGCATCAGGGGCTCGCGCAGCCGAACTGTTGCCAGGC 4246
4351 ATCAGGATGATCTGCAAGAGAGCATCAGGGGCTCGCGCAGCCGAACTGTTGCCAGGC 4410
4247 TCAAGCGCGCATGCCCGAGCGGAGGATCTGCTGTGACCCATGCGGATGCGCTGCTTC 4306
4411 TCAAGCGCGCATGCCCGAGCGGAGGATCTGCTGTGACCCATGCGGATGCGCTGCTTC 4470
4307 CGAATATCATGGTGGAAATGCGCGCTTTCTGATTCATCGACTGCGCGGCTGGTG 4366
4471 CGAATATCATGGTGGAAATGCGCGCTTTCTGATTCATCGACTGCGCGGCTGGTG 4530
4367 TGGCGGACCGCTATCAGACATACGTTGGCTACCGCTGATATGCTGAAGAGCTTGGCG 4426
4531 TGGCGGACCGCTATCAGACATACGTTGGCTACCGCTGATATGCTGAAGAGCTTGGCG 4590
4427 GCGAATGGCTGACCGCTTCTCTGTCTTTACGCTATCGCGCTCCCGATTCGCGAGCGCA 4486
4591 GCGAATGGCTGACCGCTTCTCTGTCTTTACGCTATCGCGCTCCCGATTCGCGAGCGCA 4650
4487 TCGCTTCTATCGCTTCTTGAAGAGCTTCTTGAAGAGCTTCTTGAAGAGCTTCTTGAAGAG 4546
4651 TCGCTTCTATCGCTTCTTGAAGAGCTTCTTGAAGAGCTTCTTGAAGAGCTTCTTGAAGAG 4710
4547 AAGAGATTTATTTAGTCTCCAGAAAGGGGGAATGAAGACCCCACTGATAGTTTG 4606
4711 AAGAGATTTATTTAGTCTCCAGAAAGGGGGAATGAAGACCCCACTGATAGTTTG 4770
4607 GCAAGCTAGCTTAAGTAACGCTTTTGAAGGATGGAATAATACATAACTGAGATAG 4666
4771 GCAAGCTAGCTTAAGTAACGCTTTTGAAGGATGGAATAATACATAACTGAGATAG 4830
4667 AGAAGTTTCAGATCAAGGTCAAGGACAGATGAGACAGCTGATATGAGGCGCAACAGGATAT 4726
4831 AGAAGTTTCAGATCAAGGTCAAGGACAGATGAGACAGCTGATATGAGGCGCAACAGGATAT 4890
4727 CTGTGGTAAGCAGTTCTGCGCGGCTCAGGGCCAGAACAGATGGAAGCTGATATG 4786
4891 CTGTGGTAAGCAGTTCTGCGCGGCTCAGGGCCAGAACAGATGGAAGCTGATATG 4950
4787 GGCCAAACAGGATATCTGTGTAGCAGTTCTCTGCCCGGCTCAGGGCCAGAACAGATG 4846
4951 GGCCAAACAGGATATCTGTGTAGCAGTTCTCTGCCCGGCTCAGGGCCAGAACAGATG 5010
4847 GTCCCCAGATCGCGTCCAGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGT 4906

5011 GTCCCCAGATCGCGTCCAGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGT 5070
4907 GCCCCAAGGACCTGAATGACCCCTGTGCTTATTTGAACCTAAACCAATCAGTTGCTTCTC 4966
5071 GCCCCAAGGACCTGAATGACCCCTGTGCTTATTTGAACCTAAACCAATCAGTTGCTTCTC 5130
4967 GCTTCTGTTCCGGGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTC 5026
5131 GCTTCTGTTCCGGGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTC 5190
5027 GGGCGCCAGTCTCCGATGACTGAGTCCCGGGTACCGGTATCCCAATAAACCCCTC 5086
5191 GGGCGCCAGTCTCCGATGACTGAGTCCCGGGTACCGGTATCCCAATAAACCCCTC 5250
5087 TTGCAGTTGATCCGACTTGTGCTGCTTCTTGGGAGGCTCTCTCTGAGTGATT 5146
5251 TTGCAGTTGATCCGACTTGTGCTGCTTCTTGGGAGGCTCTCTCTGAGTGATT 5310
5147 GACTACCCGTACGCGGGGCTTTTCATTGG 5177
5311 GACTACCCGTACGCGGGGCTTTTCATTGG 5341

RESULT 4
BD136838
LOCUS
DEFINITION
BD136838
Transsonomics associated with gene transfer into mammary epithelial cells.
ACCESSION
BD136838.1 GI:23231783
VERSION
JP 2002504365-A/1.
KEYWORDS
unidentified
SOURCE
unidentified
ORGANISM
unclassified.
REFERENCE
1 (bases 1 to 7699)
AUTHORS
Gavora,J.S., Falconer,M.M., Nguyen,T.H. and Benkel,B.P.
TITLE
Transsonomics associated with gene transfer into mammary epithelial cells
JOURNAL
Patent: JP 2002504365-A 1 12-FEB-2002;
HER MAJESTY IN RIGHT OF CANADA AS REPRESENTED BY THE MINISTER OF AGRICULTURE AND AGRI FOOD CANADA
COMMENT
OS Unidentified
PN JP 2002504365-A/1
PD 12-FEB-2002
PF 25-JUN-1998 JP 2000533535
PR 24-FEB-1998 CA 2224108
PI JAN S GAVORA,MARCIA M FALCONER,THUY H NGUYEN,BERNHARD F BENKEL
PC A01K67/027,C12N9/64,C12N15/09,C12P21/00,C12N15/00 CC
Strandedness: Double;
CC Topology: Circular;
CC Transsonomics associated with gene transfer into mammary CC epithelial cells

FEATURES
source
FH Key Location/Qualifiers
FT source 1..7699
FT /organism='Unidentified'.
FT Location/Qualifiers
1..7699
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN
Query Match 40.2%; Score 2078.8; DB 6; Length 7699;
Best Local Similarity 64.3%; Pred. No. 0;
Matches 3339; Conservative 0; Mismatches 1814; Indels 38; Gaps 14;

QY 1 TTTTGAAGAGCCCCCAGCTAGGTGGCAAGCTAGCTTAAGTAACGCCACTTTGCAAGGCAT 60
DB 175 TTTTGAAGAGCCCCCAGCTAGGTGGCAAGCTAGCTTAAGTAACGCCACTTTGCAAGGCAT 234
QY 61 GGAAAAATACATACTGAGAAATAGAAAAGTTTCAGATCAAGGTCAGGAACAAAGAACAGC 120

235 GGAATAATACATACTGAGATAGAAAAAGTTTCAGATCAAGGTTCAGGAACAAAGAAACAGC 294
121 TGAATACCAACAGGATATCTGTGTAGCGGTTCCTGCCCGGCTCAGGCCCAAGAAC 180
295 TGAATACCAACAGGATATCTGTGTAGCGGTTCCTGCCCGGCTCAGGCCCAAGAAC 354
181 GATGAGACAGCTGAGTGTATGAGGCGCAACAGGATATCTGTGTAGCGGTTCCTGCCCGG 240
355 GATGAGACAGCTGAGTGTATGAGGCGCAACAGGATATCTGTGTAGCGGTTCCTGCCCGG 414
241 CTGGGGCAAGAACAGATGTCTCCAGATCGGTTCAGCCCTCAGCAGTTTCTAGTGAA 300
415 CTGGGGCAAGAACAGATGTCTCCAGATCGGTTCAGCCCTCAGCAGTTTCTAGTGAA 474
301 TCATCAGATGTTCCAGGTCGCCCAAGGACCTGAAATGACCTGTACCTTATTTGAAC 360
475 TCATCAGATGTTCCAGGTCGCCCAAGGACCTGAAATGACCTGTACCTTATTTGAAC 534
361 TAACCAATCAGTTCCGCTTCTGCTTCTGTTCCGCGCTTCCGCTCTCCGAGCTCAATAA 420
535 TAACCAATCAGTTCCGCTTCTGTTCCGCGCTTCCGCTCTCCGAGCTCAATAA 594
421 AGAGCCCAACCCCTCACTCGGCGGCGCAGTCTTCGATAGACTGCGTCCCGGGTAC 480
595 AGAGCCCAACCCCTCACTCGGCGGCGCAGTCTTCGATAGACTGCGTCCCGGGTAC 654
481 CCGTATTTCCCAATAAGCCTCTGCTGTTTCATCCGATCCGATCGTGTCTGCTGTTCTTG 540
655 CCGTATTTCCCAATAAGCCTCTGCTGTTTCATCCGATCCGATCGTGTCTGCTGTTCTTG 714
541 GGAGGGTCTCTCTGAGTGTGATCTACCCAGACGGGGTCTTTCATTTGGGGGCTCGT 600
715 GGAGGGTCTCTCTGAGTGTGATCTACCCAGACGGGGTCTTTCATTTGGGGGCTCGT 774
601 CCGGATTTGAGACCCCTGCCCCAGGACACCGACCCACCGAGGTAAGCTGGCC 660
775 CCGGATTTGAGACCCCTGCCCCAGGACACCGACCCACCGAGGTAAGCTGGCC 834
661 AGCAACTTATCTGTCTGCTCGGATGCTCTAGTCTCTATGTTGATGTTATGCGCTGG 720
835 AGCAACTTATCTGTCTGCTCGGATGCTCTAGTCTCTATGTTGATGTTATGCGCTGG 894
721 TCTGTACTAGTTAGTAACTAGCTCTGTATCTGCGGACCCGCTGGTGAACCTGACGAGTT 780
895 TCTGTACTAGTTAGTAACTAGCTCTGTATCTGCGGACCCGCTGGTGAACCTGACGAGTT 954
781 CTGAACACCCCGCGCAACCCCTGGGAGACGTCCTCAGGACCTTTGGGGCGCTTTTGTG 840
955 CTGAACACCCCGCGCAACCCCTGGGAGACGTCCTCAGGACCTTTGGGGCGCTTTTGTG 1014
841 CCGACCTGAGGAAGGAGTCGATGTGGAATCCGACCCCGTCAGGATATGTTCTGTT 900
1015 CCGACCTGAGGAAGGAGTCGATGTGGAATCCGACCCCGTCAGGATATGTTCTGTT 1074
901 AGGAGACGAGAACCTAAACAGTTTCCCGCTTCTGAAATTTTGTCTTCCGTTTGGAA 960
1075 AGGAGACGAGAACCTAAACAGTTTCCCGCTTCTGAAATTTTGTCTTCCGTTTGGAA 1134
961 CCGAAGCGCGCGCTTGTCTGCTGACGACCAAGCTTGGGTCGAGGTCGACTCTAGAGGA 1020
1135 CCGAAGCGCGCGCTTGTCTGCTGACGCGCTGAGCATCGTCTGTTGTTCTCTGCT 1194
1021 TCAATTCGGACGAGTAATCGGTGCTGCTCTTTAGGACATATGAAAGTATGACAGT 1080
1195 GACTGTGTTTCTGTATTTGCTGAAAAATTAGGGCCAGATGTTTACCCTCCCTTAAGTTT 1254
1081 GGGATGACTTCTCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTCG 1140
1255 GACCTTAGTCACTGGAAAGATGTGACAGCGATCGCTCAACACCGCTGGTAGATGCAA 1314
1141 ATGCTCGCTCAGTGACAGCTTTGCTTCTCCCACTCTCAATAAATGGCCCAACTCTTCAAG 1200
1315 GAAGAGAGCTTTGGGTACCTTCTGCTCTGAGATGCGCAACCTTTAAAGCTGCGATGGCC 1374

QY 1201 AGAGGTGAAGTCTTATAAAACATGCACTGATCACCTTTTATCTCATTTGTGTTTCTAGTTC 1260
Db 1375 GCGAGACGGCACCTT-TAACCGAGACCTCATCACCCAGGTTAAGATCAAGGTCTT-TTC 1431
QY 1261 TCGTGCCCATCATTTGGCATAGTGGCAGCTCAGCTCTCTGAAATGGGAAACGAAGAAATTGCA 1320
Db 1432 ACCTGGCCGATGGACACCCAGACAGGTCCCCTACATCGTACCTGGGAAGCTTTGGC 1491
QY 1321 CGGTTGGCTCAGTTAATGACAGATATATCTCAAGTCCGGAGGCAAAAGGAAATGGCAGTG 1380
Db 1492 TTTTGAACCCCTCCCTGGGTCAAGCCCT-----TTGTACACCCCTAAGACCTCCGCCCTCT 1546
QY 1381 AAGATGAAATGAGATTTCCGAGAAAGCTGTGTATGGAACGATGAGCAACATCGAAAGCAGAA 1440
Db 1547 CTTCTCTCAATCCGCCCGCTCTCTCCCTTTGAACCTCTCTGTTGACCCCGCCCTCGATCC 1606
QY 1441 TCAGATATCTTTTCAGATAATGAAGCCAAATCTCTAGATGCTAAGAAATTTCCAAAAATTCA 1500
Db 1607 TCCCTTTATCCAGCCCTCACTCTCTCTAGGCGCGGAATTCGTTAACTCGACATGGAA 1666
QY 1501 GCATAACACTGATCAAGATTTAATGATGTTCTTTTCCAGCTAAATTCCTTACTTTCCT 1560
Db 1667 GTCCCTTCTCTCTCGCAGCTGTGGGCTTTGCTGGGCACAGTACAAATCCCAACACTCAG 1726
QY 1561 CCATCCAGG--AACATGAGAAATATCATAGGGGATATCTCCAAGTCAATTAGTAGTCTGAA 1618
Db 1727 GCTGGGAGGACATCTATCTGTCGATCTCTTTGAATGGCGCTGGGCCGACATTTGCACTGAG 1786
QY 1619 CACCACAGTACTTTGATTGAGTTTCAGTATTTGAACACTGAATGGCAGAGTCCCAAGAGAA 1678
Db 1787 TCGGAACACTATTTAGCTCTCTAATGGTTTGGAGGAGTTTCAGGTTTCTCTCCCAATGAA 1846
QY 1679 TGATTTAAACAACAGAGAGAGATGCGTAAATTTAGAGGCGGTATATACAAATGATCAGC 1738
Db 1847 AACATTTGCTACTTAATCCGAAACAGGCCCTGGTGGGAAAGATACCAAGCCCATCAGCTAC 1906
QY 1739 AGAAATTTAAGTCTCTAGATGAAACAAAGTATATTTTGAACACAGGAAATAAAGGGGAAAT 1798
Db 1907 AAGATCTGAGTCGATCGGGCAATGAAATGAATTCAGAGACATGGTGACAGATGCAAC 1966
QY 1799 GAAACTGTTGAATAATATCACTAATGATCTGAGGCTGAAAGGATTTGGGAACATTTCTAGAC 1858
Db 1967 AATGTTGGAG-TTCGTTATTTATGTGGATGCTGTTGTCAATCACATGTGTGATCTATGGG 2025
QY 1859 ATTGAAAAATATCACTTTACTCCAAGTGCAGAAAGTGTCTCGTGACTGGGAAATGGAC 1918
Db 2026 TGGCACGGGCACCCACTCAACATGTGGGAGCTATTTCAACACCCGGGACTAGAGATTTTCC 2085
QY 1919 CAACGATCTGGGCTCCAACATGACCATCGGGGCTGTGAACAGACAGAGGTGAATTCACAGG 1978
Db 2086 CGCTGTGCGGTACTCTGCTCGGGATTTCAATGACGGCAATGTACACTGCAAGTGGAGA 2145
QY 1979 CACCTATACACAGCCGTAAACAGCCACATCAAAATGAGATCAAAAGAGTCAACACTGCAATGG 2038
Db 2146 CATGAAAAATTATGGGACATGTATCAGSTCCGGAATTGCAAGTTGTCCAGCTTCTTGA 2205
QY 2039 GACA-CAAAACACCATCAACAGAGGACCCAGCCACCTTTGGCTTCACCGTCAATTTGGA 2097
Db 2206 TCTGGCTCTGAGAAAGACTATGTACGCTCAAAATTTGACGCTACATGAATCACTCAT 2265
QY 2098 AGTTTTCAGAGTCCACCATGCTTTCACGGGCGAGTGTCTCATAGACAGGAATGGGAAGG 2157
Db 2266 TGATATGGGTGTAGCGGGTTCCGGATCGATGCTGCCACGATATGTGGCCAGGGGACAT 2325
QY 2158 AGTCTCTGAAGACCATGTGGCTGCTGGGTCAAGTGTTAATGACATTTGGTGTGATGACTGGA 2217
Db 2326 AAGAGCGTTTCTGGAACAACTGCACCGATCTAAATACTCACTAGTGGTTTTCAGCAGGAACGAA 2385
QY 2218 AAGCTACAGGGTC-GGCATCAACATCTTCACTGCGCTCGGCAACAGAAAGGAGTAGTG 2276
Db 2386 ACCCTTTATTTACCAAGAGGTAATTGACTTGGGAGGAGGCCAATCAACAGCAGTCACTA 2445

QY 2277 AGTGACCAAGTCTCTCTG---GACTCCAGGTGAAGAAGAGATAGAGGCCCTCTCTGGAC 2333
Db 2446 CTTTGGGATGGCCGCGTGAAGAAATCAAGATAGTGTCGCAAACTGGGGAGCGGTGATCCG 2505
QY 2334 AAAATGGTATACCAAGGCTTTCAGGTCTAAATAGGTACTCCAGGTCTTAAAGGTGATCGGG 2393
Db 2506 GAAGTGAATGGAGAGAAGATGGCTACTTAAAGAACTGGGGAGAGGCTGGGCTTTGT 2565
QY 2394 GGGATCTCTGTTTACCTGGAGTTCGAGGATTCAGGAC--CAATGGGGAAGACCGGGA 2451
Db 2566 GCCTTCTGACAGAGCCCTGGTGTGTGTGTGAATAACCAACGACCAACCGGGGGGACCGGGC 2625
QY 2452 AGCCAGGACTTAAATGACAAAGGCCAGAGGAGGAAGAAGGAGTGGAGCATGCNAA 2511
Db 2626 AGGCGAGCTTCCATTTTCTTCTGGATGCGCAGGCTTTATANAATGGCGGTGGTTT 2685
QY 2512 GACAACTTAATACAGTCCGACTGGTGGGTGGCAGCGGCCCTCAGCAAGGAGAGTGGAGA 2571
Db 2686 CATGCTGCTCATCGTACGGGTTTACACCGGGTGTATGTCAGTTTATCGTTGGCCAAAGTA 2745
QY 2572 TTTTTCAGAAAGGCGAGTGGGTTAGGTGTGTGACGACCGCTGGGAACCTGGTGGAGAC 2631
Db 2746 TTTCCAAAACGAGTGGATGTTAAACGACTGGGTGGGACCAACCAAGTAACTCGGACGGATC 2805
QY 2632 T-----GGTCTCTGACGAGCTTGGGATACAAAGGTGTTCAAAGTGTGCATAAGCGAGCT 2687
Db 2806 GACGAATCCGTTTCAATCAATCAACGAGACACTACCTGTGGCAATGACTGGGTCTCGGAACA 2865
QY 2688 TATTTGGAAAGGTACGGGTCCAAATATGGCTGAATGAATGATTTTGGTTCGGGAAAGAG 2747
Db 2866 TCGCTGGGCAAAATAGG-----ACATGGTTATCTCCGTAATGTGGTAGACGGTCAGC 2921
QY 2748 TCATPCCATTAAGAGTGCAGAAATPAGACAGTGGGGTGTGAGAGCTGTTCGACAGCGAA 2807
Db 2922 CTTTCTCAAACTGGTGGACAAACGGGAGCAATCAAGTAGCTTTCCGTGGGGGACAGAG 2981
QY 2808 GATCTGGGGTCACTTGGACCTACATATGCAATATTTTCAATTCATTCACATTTTTTAAA 2867
Db 2982 GCTTCATTTGTTTAAATAGTACTGGTATATGAATGCTGATTTGCAAACTGGTCTGC 3041
QY 2868 CTGTTATAAGTGAATTTTTTCTTCTTCTTCACTAAATCAGCTTAATTAATATTAAGA 2927
Db 3042 CTGCTGNAACCTACTCGATGTTTATTTCTGGACAAAAGGAAGGAGTGGGTGACTGGAA 3101
QY 2928 AACTAAGAAATTTTATCCAGAAAAGGAATATTTAAAAATCACTGGATAAACAATATAAA 2987
Db 3102 AGCAGGTGTACGTTCTCGGATGAAAGGCCAATTTCCAGATTAGTAAACAGCGATGAAG 3161
QY 2988 TAGCTTCATATTTGCTTCAAAATACAGAACCAATTTCACTTCTCTAGGTTTTTAAGTGGC 3047
Db 3162 ATCCATTTGTTGCAATTCACGTTGATGCCAAGTTTAAAGCTTCGAGGATCCACTAGTAAC 3221
QY 3048 TCGTGGCGAATGATCCCTCAGGATATAGTATAGTTTCGCTTTTGCATAGGAGGGGAAA 3107
Db 3222 GGCGCGAGTGTCTGGAATTCGG-----CTGTGCAATCTAGGGCGGCAATTCGCGC 3276
QY 3108 TGTAAGTCTTATGCAATCTCTTGTAGTCTTGCAACATGTTAAACCATGATGATAGCAATG 3167
Db 3277 CCTCTCCCCCCCCCTTAAGTTACTGGCGAAGCGCTTGGATTAAGGCCGTGTGTG 3336
QY 3168 CTTTAAAGGAGAAAAGAACCGGTGATCCGATTTGGTGGAAAGTAAAGTGTGTACGATC 3227
Db 3337 TTTGCTATATGTTGTTTCCACCATATTTCCCATATTTGCCGTCTTTTGGCA-----TGTGAGGGCC 3390
QY 3228 GTGCTTTATAGGAAGGCAACAGACGGGTCTGACATGGATTTGGACGAACCACTGAATTC 3287
Db 3391 CGSAAACCTGGCCCTGTCTTCTGACGAGCAATTCCTAGGGGTCTTTCCCTCTCGCCAAA 3450
QY 3288 GCATTTGACAGATATTTGATTTAAGTGTCTAGCTCGATACAGCAAAACCGCATTTTGACCAT 3347
Db 3451 GGAATGCAGGTCTGTGTAATGTCGTGAGGAGCAGTTCCTCTGGAAGCTTCTTTGAAGA 3510
QY 3348 TCACCAATTGGTGTGCACTCTCAAGCTTTCAGCTGCCGCAAGCACTCAGGGCGCAAGGG 3407

Db 3511 CAACAAACGTCTGTAGGCAACCTTTTGAGGAGCGGAACCCCCACCTCTGGCGACAGGTGC 3570
QY 3408 CTGCTAAAGGAAGCGGAACACACGTAGAAAGCGAGTCCGACAGAAAACGCTGCTGACCCCGAT 3467
Db 3571 CTCTGGCGCCAAAAGCCACCGTGTATAGATACACCTGCAAAAGCGGCACAAACCCAGTGC 3630
QY 3468 GAATGTAGCTACTGGGCTATCTGGACA-AGGNAACGACGAGCGCAAGGAAGACAGG 3526
Db 3631 CACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAATGGCTCTCTCAAGCGTAGTCAAC 3690
QY 3527 TAGCTTCAGTGGGCTTTACATGGCATAGTCTAGATCGGCGGTTTTATGGACACAAGCG 3586
Db 3691 AAGGGCTGAAGGATGCCCAAGAGTACCCCATTTGTATGGGAATCTGATCTGGGGCCTCG 3750
QY 3587 AACCGGAATTTCCAGCTGGGGCGCCCTCTGCTGAAGTTGGGAGCCCTGCAAAAGTAAACT 3646
Db 3751 GTGCACATGCTTTACATGTTGTTTAGTCGAGTTTAAAAAAGCTCTPAGGCCCCCGAACAC 3810
QY 3647 GGATGGCTTTCTTGGCCCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGA 3706
Db 3811 GGGACGCTGGTTTTCTTTGAAAAACACGATGATAGCTTTGCCACAACCCCGGATAATTC 3870
QY 3707 CAGATGAGGATCGTTTCGCATGATTCAAACAAGATGATTCGACGAGGTTCTCCGGCG 3766
Db 3871 CTGCAGCAATATGGGATCGGCCAATTGAACAAGATGATTGCACGCAAGTTCTCCGGCG 3930
QY 3767 CTTGGGTGGAGAGGCTATTCCGGCTATGACTGGGCAACAAGACAACTCGGCTCTGATG 3826
Db 3931 CTTGGGTGGAGAGGCTATTCCGGCTATGACTGGGCAACAAGACAACTCGGCTCTGATG 3990
QY 3827 CCSCCGTGTTCGGCTGTGACGCGAGGGCGCCGGTCTTTTGTCAAGACCACTGT 3886
Db 3991 CCGCCGTTTCCGGCTGTGACGCGAGGGCGCCGGTCTCTTTTGTCAAGACCACTGT 4050
QY 3887 CCGTGCCTTGAATGAATGACTGACGAGAGGCGGGCTACTCGTGGCTGGCCAGGAGCG 3946
Db 4051 CCGTGCCTTGAATGAATGACTGACGAGAGGCGGGCTACTCGTGGCTGGCCAGGAGCG 4110
QY 3947 GCGTTCCTTCGCGAGCTGTCTCGACGTTGCTCACTGAAGCGGAGGAGACTGGCTGTAT 4006
Db 4111 GCGTTCCTTCGCGAGCTGTCTCGACGTTGCTCACTGAAGCGGAGGAGACTGGCTGTAT 4170
QY 4007 TGGCGAAGTTCGGGGCAGGATCTCTGTCTCATCTCACTTGTCTCTGCGCAGAAAGTAT 4066
Db 4171 TGGCGAAGTTCGGGGCAGGATCTCTGTCTCATCTCACTTGTCTCTGCGCAGAAAGTAT 4230
QY 4067 CCAATGCTGTGATGCAATCGGCGGCTGCATAGCTTGTGATCGGCTACCTGCGCATTCG 4126
Db 4231 CCAATGCTGTGATGCAATCGGCGGCTGCATAGCTTGTGATCGGCTACCTGCGCATTCG 4290
QY 4127 ACCAACAGGAAACATCGCATCGAGGAGCAGTACTCGGATGGAAGCGGCTCTTGTG 4186
Db 4291 ACCAACAGGAAACATCGCATCGAGGAGCAGTACTCGGATGGAAGCGGCTCTTGTG 4350
QY 4187 ATCAGATGATCTGGAGAAAGACATCAGGGGCTCGCGCCAGCGCAACTGTTCGCCAGGC 4246
Db 4351 ATCAGATGATCTGGAGAAAGACATCAGGGGCTCGCGCCAGCGCAACTGTTCGCCAGGC 4410
QY 4247 TCAAGGGCGCATGCGGAGCGGAGGATCTCGTGTGACCCATGGCGATGCTCTGCTTCG 4306
Db 4411 TCAAGGGCGCATGCGGAGCGGAGGATCTCGTGTGACCCATGGCGATGCTCTGCTTCG 4470
QY 4307 CGAATATCATGTGGAAAAATGGCCGCTTTCTGATTCATCGATGTGGCGCGCTGGGTG 4366
Db 4471 CGAATATCATGTGGAAAAATGGCCGCTTTCTGATTCATCGATGTGGCGCGCTGGGTG 4530
QY 4367 TGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTTGCTGAAGAGCTTCGCG 4426
Db 4531 TGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTTGCTGAAGAGCTTCGCG 4590
QY 4427 GCBAATGGGCTGACCGCTTCTCTGTCTTTACGGTATCGGCTTCGCGCTCCGATTCGAGCGCA 4486

Db 4591 GCGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTCGCGAGCGCA 4650
Qy TCAGCTTCTATCGCTTCTTCCAGAGTTCCTTCCAGCGGAGTCTCGGGTTCGATATAAAT 4546
Db TCAGCTTCTATCGCTTCTTCCAGAGTTCCTTCCAGCGGAGTCTCGGGTTCGATATAAAT 4710
Qy 4547 AAAAGATTTTATTTAGTCTCCAGAAAAAGGGGGGAAATGAAAGACCCACCTGTAGGTTTG 4606
Db 4711 AAAAGATTTTATTTAGTCTCCAGAAAAAGGGGGGAAATGAAAGACCCACCTGTAGGTTTG 4770
Qy 4607 GCAAGCTAGCTTAAGTAAAGCCATTTTGCAGAGGCAATGGAAGAAATACATAACTGAGATAG 4666
Db 4771 GCAAGCTAGCTTAAGTAAAGCCATTTTGCAGAGGCAATGGAAGAAATACATAACTGAGATAG 4830
Qy 4667 AGAAGTTCAGATCAAGGTCAGGACAGATGAAACAGCTGAAATATGGGCCAACAGGATAT 4726
Db 4831 AGAAGTTCAGATCAAGGTCAGGACAGATGAAACAGCTGAAATATGGGCCAACAGGATAT 4890
Qy 4727 CTGTGGTAAGCAGTTCTGCCCCCGCTCAGGGCCAAAGAACAGATGGAACAGCTGAATATG 4786
Db 4891 CTGTGGTAAGCAGTTCTGCCCCCGCTCAGGGCCAAAGAACAGATGGAACAGCTGAATATG 4950
Qy 4787 GGCACAAACAGATATCTGTGTAAAGCAGTTCCTGCCCCCGCTCAGGGCCAAAGAACAGATG 4846
Db 4951 GGCACAAACAGATATCTGTGTAAAGCAGTTCCTGCCCCCGCTCAGGGCCAAAGAACAGATG 5010
Qy 4847 GTCCCCAGATGGGTCAGGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGT 4906
Db 5011 GTCCCCAGATGGGTCAGGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGT 5070
Qy 4907 GCCCCAAGGACCTGAAATGACCTGTGCTTATTTGAACTAAACCAATCAGTTCCGTTCTC 4966
Db 5071 GCCCCAAGGACCTGAAATGACCTGTGCTTATTTGAACTAAACCAATCAGTTCCGTTCTC 5130
Qy 4967 GCTTCTGTTCGGCGCTTCTGCTCCGAGCTCAATAAAGAGCCCAACACCCCTCACTC 5026
Db 5131 GCTTCTGTTCGGCGCTTCTGCTCCGAGCTCAATAAAGAGCCCAACACCCCTCACTC 5190
Qy 5027 GGGGGCCAGTCCCGATTGACTGAGTCGCGCGGGTACCCGTGTATCCATTAACCCCTC 5086
Db 5191 GGGGGCCAGTCCCGATTGACTGAGTCGCGCGGGTACCCGTGTATCCATTAACCCCTC 5250
Qy 5087 TTGCAGTTGCAATCCGACTTGTGGTCTCGCTGTCTTGGAGGGTCTCCTCTGAGTGATT 5146
Db 5251 TTGCAGTTGCAATCCGACTTGTGGTCTCGCTGTCTTGGAGGGTCTCCTCTGAGTGATT 5310
Qy 5147 GACTACCCGTGAGGGGGTCTTTCAATTTGG 5177
Db 5311 GACTACCCGTGAGGGGGTCTTTCAATTTGG 5341

RESULT 5
AR381629
LOCUS AR381629 7699 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6608238.
ACCESSION AR381629
VERSION AR381629.1 GI:40089782
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 7699)
AUTHORS Gavora,J.S.; Falconer,M.M.; Nguyen,T.H. and Benkel,B.P.
TITLE Trans-somatics with gene transfer into mammary epithelial cells
JOURNAL Patent: US 6608238-A 1 19-AUG-2003;
Gala Design Inc.; Middleton, WI;
CAX;
FEATURES
source Location/Qualifiers
1..7699
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Query Match 40.2%; Score 2078.8; DB 6; Length 7699;
Best Local Similarity 64.3%; Pred. No. 0;
Matches 3339; Conservative 0; Mismatches 1814; Indels 38; Gaps 14;
Qy 1 TTTGAAAGACCCACCCGTAGGTGGCAAGCTTAAGTAAACGCCACTTTTGCAGGCAT 60
Db TTTGAAAGACCCACCCGTAGGTGGCAAGCTTAAGTAAACGCCACTTTTGCAGGCAT 234
Qy 61 GGAATAATACATAACTAGAAATAGAAAAGTTTCAGATCAAGSTCAGGACCAAGAAACAGC 120
Db 235 GGAATAATACATAACTAGAAATAGAAAAGTTTCAGATCAAGSTCAGGACCAAGAAACAGC 294
Qy 121 TGAATACCAACAGGATATCTGTGGTAAAGCGTTTCTGCCCGGCTCAGGGCCCAAGAAC 180
Db 295 TGAATACCAACAGGATATCTGTGGTAAAGCGTTTCTGCCCGGCTCAGGGCCCAAGAAC 354
Qy 181 GATGAGACAGCTGAGTGATGGGCCAAACAGGATATCTGTGTAAAGCAGTTTCTGCCCGG 240
Db 355 GATGAGACAGCTGAGTGATGGGCCAAACAGGATATCTGTGTAAAGCAGTTTCTGCCCGG 414
Qy 241 CTGGGGCCAAAGACAGATGGTCCCGAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGAA 300
Db 415 CTGGGGCCAAAGACAGATGGTCCCGAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGAA 474
Qy 301 TCATCAGATGTTTCCAGGGTCCCGAGGACCTGAAATGACCCCTGACCTTATTTGAAC 360
Db 475 TCATCAGATGTTTCCAGGGTCCCGAGGACCTGAAATGACCCCTGACCTTATTTGAAC 534
Qy 361 TAACCAATCAGTTTCGCTTCTGCTTCTGTTGGCGGCTTCCGCTCTCCGAGCTCAATAAA 420
Db 535 TAACCAATCAGTTTCGCTTCTGCTTCTGTTGGCGGCTTCCGCTCTCCGAGCTCAATAAA 594
Qy 421 AGAGCCACAAACCCCTCAGTCCGGCGCCAGTCTTCCGATAGACTGGTCCCGGGTAC 480
Db 595 AGAGCCACAAACCCCTCAGTCCGGCGCCAGTCTTCCGATAGACTGGTCCCGGGTAC 654
Qy 481 CCGTATTTCCCAATAAAGCCTCTTGTCTTTGCAATCCGAATCGTGGTCTCGCTGTTCCTTG 540
Db 655 CCGTATTTCCCAATAAAGCCTCTTGTCTTTGCAATCCGAATCGTGGTCTCGCTGTTCCTTG 714
Qy 541 GGAAGGCTCTCTCTGAGTGATGACTACCAAGCGGGGCTTTTCAATTTGGGGGCTCGT 600
Db 715 GGAAGGCTCTCTCTGAGTGATGACTACCAAGCGGGGCTTTTCAATTTGGGGGCTCGT 774
Qy 601 CCGGATTTGAGACCCCTGCGGAGACCAAGGACCAAGGACCAAGGAGGTAAGCTGGCC 660
Db 775 CCGGATTTGAGACCCCTGCGGAGACCAAGGACCAAGGAGGTAAGCTGGCC 834
Qy 661 AGCAACTTATCTGTCTGTCCGATTTGCTAGTGTCTATGTTTGAATGTTATGCGCCTGCG 720
Db 835 AGCAACTTATCTGTCTGTCCGATTTGCTAGTGTCTATGTTTGAATGTTATGCGCCTGCG 894
Qy 721 TCTGTACTAGTTAGTAACTAGCTCTGTATCTGGCGGACCCGTGGTGGTGAACGAGATT 780
Db 895 TCTGTACTAGTTAGTAACTAGCTCTGTATCTGGCGGACCCGTGGTGGTGAACGAGATT 954
Qy 781 CTGAACACCCCGCGGCAACCTCGGAGACGCTCCAGGAGCTTTGGGGCCCGTTTGTGG 840
Db 955 CTGAACACCCCGCGGCAACCTCGGAGACGCTCCAGGAGCTTTGGGGCCCGTTTGTGG 1014
Qy 841 CCGACCTGAGGAGGAGTGCATGTGGAATCCGACCCCGTCAGGATATGTGGTCTCGT 900
Db 1015 CCGACCTGAGGAGGAGTGCATGTGGAATCCGACCCCGTCAGGATATGTGGTCTCGT 1074
Qy 901 AGGAGACGAGAACCTAAACAGTTCCCGCTCCGCTCGAATTTTGTCTTTCGGTTTGAA 960
Db 1075 AGGAGACGAGAACCTAAACAGTTCCCGCTCCGCTCGAATTTTGTCTTTCGGTTTGAA 1134
Qy 961 CCGAGCCCGCGCTCTTGTCTGTGAGCCAGCTTGGGCTGACGCTCGACTAGAGGA 1020
Db 1135 CCGAGCCCGCGCTCTTGTCTGTGAGCCGCTGACGATCGTCTGTGTGTCTGTCT 1194
Qy 1021 TCAATTCGGCACGAGTAAATCGGTGCTGCGGTCTTTAGGACATATGAAGATGTCACAGT 1080

Db 1195 GACTGTGTTCTGTAATTTGCTGAAATTAAGGCGCAGACTGTTACCACTCCCTTAAAGTTT 1254
Qy 1081 GGGATGACTTTCTCTGATCAGCAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTGG 1140
Db 1255 GACCTTAGTCTCACTGGAAAGATGTCGAGCGGATCGCTCACAAACAGTCGGTAGATGCAA 1314
Qy 1141 ATGCTCGCTCAGTGACAGCTTTGCTTCTCCCATCTTAATAAATGGCCCAACTCTTCAAG 1200
Db 1315 GAAGAGACGTTGGGTACCTTCTGCTGCAAGATGGCCAACTTTAAACGTCGATGGCC 1374
Qy 1201 AGAGGATGAAGTCTTATAAAATCAGCTGATCACCTTTTATCTCATTTGTTCTAGTTTC 1260
Db 1375 GCGAGACGGCACCTT-TAACCGAGACCTTCATCACCGAGTTTAAGATCAAGGTCTT-TTC 1431
Qy 1261 TCGTGCCCATCATTTGGCATAGTGGCAGCTCAGCTCCCTGAATGGGAACCAAGAAATGGCA 1320
Db 1432 ACCTGGCCCGCATGACACACCCAGACAGGTCCTTACATCGTGACCTGGGAAGCCTTGGC 1491
Qy 1321 CGGTGGCTCAGTTAATGCGAGATATATCTCAAGTCCGGGAAGCAAGAAATGGCAGTG 1380
Db 1492 TTTTGACCCCTCCCTCGGTCAAGCCCT-----TTGTACACCCCTAAGCCCTCCGCCCTCT 1546
Qy 1381 AAGATGAATGAGATTTCCGAGAGCTGTGAGAAACGATGAGCAACATGGAAGCAGAA 1440
Db 1547 CTTCCCTCATCCGCCCGCTCTCTCCCGCTTGAACTCTCTGTTCCGACCCCGCTTCGATCC 1606
Qy 1441 TCCAGTATCTTTTCAAGATAATGAAGCCAACTCTCTAGATGCTTAAGAAATTTCCAAAAATTTCA 1500
Db 1607 TCCCTTTATCCAGCCCTCACTCTCTCTAGGCGCGGGAATTCGTTAACTCGACATGGAA 1666
Qy 1501 GCATAACAACTGATCAAGATTTAATGATGTTCTTTTCCAGCTAAATTCCTTACTTTCTCT 1560
Db 1667 GTCTCTCTCTCGCAGCTGTGCGGCTTTGCTGGGCACAGTACAATCCCAACACTCAG 1726
Qy 1561 CCATCCAGG--AACATGAGATATCATAGGGATATCTCCAAGTCAATTTAGTAGTCTGAA 1618
Db 1727 GCTGGGAGGACATATATCGTGCATCTCTTTGAATGGCGCTGGCGGACATTTGCACTGGAG 1786
Qy 1619 CACCACAGTACTTGATTTGAGTTTCAGTATTGAACAACACTGAATGGCAGAGTCCAAGAGAA 1678
Db 1787 TGGMACACTATTTAGCTCTTAATGGGTTTGGAGAGTTTCAGGTTTCTCTCCAAATGAA 1846
Qy 1679 TGCATTTAAACAACAGAGGAGATGCGTAATTAAGAGAGCGTATATACATGATCATCAGC 1738
Db 1847 AACATTGTCTAATCTAATCGAACAGGCCCTGGTGGGAAGATACCAGCCCATCAGCTAC 1906
Qy 1739 AGAATTAAGTCTCTAGATGAATAAACAAGTATATTTGGACAGGAATAAAGGGGAAT 1798
Db 1907 AAGATCTGCAATCGATCGGCGCAATGAATAATTCAGAGACATGSGTGACCCAGATGCAAC 1966
Qy 1799 GAAACTGTTGAATAATATCAATATGATCTGAGGCTGAAGGATGGGAACATTTCTCAGAC 1858
Db 1967 AATGTTGAG--TTGATTTAATGAGGATGCTGTTGTCATCATATGTTGGATCTATGGG 2025
Qy 1859 ATTGAAAAATATCACTTTACTCCAAAGTGGCCAGAAAGTGTCTGCTGACTGGGAATGGAC 1918
Db 2026 TGGCAGCGGCACCCACTCAACATGTGGGAGCTATTTCAACACCGGACTAGAGATTTTCC 2085
Qy 1919 CAACGATCTGGGCTCCAAACATGACCATCGGGCTGTGAAACAGCAGAGGTGAATTTCAAGG 1978
Db 2086 CGCTGTGCGTACTCTGCTGGGATTTCAATGACGGCAATGTCACTGCAAGTGGAGA 2145
Qy 1979 CACCTACATCACGCCGTAAACAGCCACATCAATAGATCAAAAGATCAAGAGTCCACCTGCATGG 2038
Db 2146 CATGAAATAATTTGGGACATGATATCAGTCCGGGATTTGCAAGTTGTCCAGCCTTCTTGA 2205
Qy 2039 GACA-CAAAAACCAATCAACAAGAGGACCCAGCCCACTTTTGGCTTCCAGCTCAATTTGGA 2097
Db 2206 TCTGGCTCTGGAGAGGACTATGTACGCTCAACAAATTTGCAGCGTACATGAATCACCTCAT 2265
Qy 2098 AGTTTTCAGAGTCCACCATGCTCTTTCACGGGCCAGTGTCTCATAGACAGGAATGGGAAGG 2157

Db 2266 TGATATGGGTGTAGCAGGGTTCCGATCGATGCTGCCAAGCATATGTGGCCAGGGGACAT 2325
Qy 2158 AGGTCTGAAAGACCATGTGCTGCTGCGGTCAAGTGTAAATGACATTTGGTGTATGACTGGA 2217
Db 2326 AAGAGCGTTTTCTGCAAAAACCTGCACGATCTAAATACTCAGTGGTTTTTTCAGCAGGAACGAA 2385
Qy 2218 AAGTACCAGGGTC-GGCATCAACATCTTCACTCGCTCGGCACACAGAGGAGTGAAGTG 2276
Db 2386 ACCCTTTTATTACCAAGAGGTAATTGACTTTGGGAGGAGAGCCAAATCAAGCAGTCAAGTA 2445
Qy 2277 AGTGACCAAGGTCTCTCTG--GACTCCAGGTGAAAAGAGAGATAGAGGGCCCTCTCTGAC 2333
Db 2446 CTTTGGGAATGGCCGCTGACAGAAATCAAGATATGGTGCCTCAACTGGGGACGGTGAATCCG 2505
Qy 2334 AAAATGGTATACCAAGGCTTTCCAGGTCTAATAGGTACTCCAGGTCTTTAAAGGTGATCGGG 2393
Db 2506 GAAATGGAATGGAGAGAAGATGGCTACTTTAAAGAACTGGGGAGAAGGCTGGGCTTTGT 2565
Qy 2394 GGGATCTCTGGTTTACTCTGAGTTCGAGGATTCAGGATTCAGGATATGGTGCCTCAACTGGGGACGGTGAATCCG 2451
Db 2566 GCCTTCTGACAGAGCCCTGGTGTGTGGATTAACACGACCAACCCAGCGGGGGGACCGGGC 2625
Qy 2452 AGCCAGGACTTAATGGACAAAAGGCCAGAGGAGGAGAAAAGGAGTGGAGATGCAAA 2511
Db 2626 AGCGGAGCTTCCATCTTACTTTCTGGGATGCCAGGCTTTATAAATGGCGGTTGGTTT 2685
Qy 2512 GACAACTAATACAGTCCGACTGGTGGTGGCAGCGCCCTCACGAAGGAGAGTGGAGA 2571
Db 2686 CATGCTCGCTCATCCGTACGGTTTACACGGGTGATGTCAAGTTTATCTGTTGGCCAAAGTA 2745
Qy 2572 TTTTTCAGGAAGCCAGTGGGTACGGTGTGTGACGACCGCTGGGAACTGGTGGAGGAC 2631
Db 2746 TTTTCGAAAACGGAGTGGATGTTAAACGACTGGGTGGGACCAACCAAGTAACTCGGACGGATC 2805
Qy 2632 T----GCTGCTCTGCGAGGCTGGGATACAAAAGTGTCAAAAGTGTGCATAAGCAGCT 2687
Db 2806 GACGAGTCCGTTACAATCAACGACAGACACTACCTGTGGCAATGACTGGGTCTCGGAACA 2865
Qy 2688 TATTTTGGAAAAGGTACGGGTCCAATATGGCTGAATGAAGTATTTTGTTCGGGAAAGAG 2747
Db 2866 TCGCTGCGACAAATAGG---AACATGTTATCTTCCGTAAATGTGTAGACGCTCAGC 2921
Qy 2748 TCATCCATTAAGAGTGCAGAAATTAAGACGTGGGGTGTGAGAGCTGTTCGCAACGAGAA 2807
Db 2922 CTTTCTCAAACTGTTGGGACAAACGGGAGCAATCAAGTAGCTTTTCGGTCGCGCGCACAGAG 2981
Qy 2808 GATGCTGGGGTCACTTTGCACTACATAATGATCATATTTTCACTTCAACATTTTAA 2867
Db 2982 GCTTCATTTGCTTTAATAATGATGAGTGTATATGAATGCGATTTGCAAACTCGTCTGC 3041
Qy 2868 CTGTTATAAGTGAATTTTTTCTTTTCTTTTCACTTAAATCAGCTTAAATTAATTTAAGA 2927
Db 3042 CTGCTGNACTACTGCGATGTTATTTCTGGACAAAGGAGGAGCGCTGCTGTAAGAA 3101
Qy 2928 AACTAAGAAATTTTATCCACAGAAAGGAATTTTAAATAATCACTGGATAAACAATATAA 2987
Db 3102 AGCAGGTGTACGTTTCTCGGATGGAAGGCCAATTTCCAGATTTAGTAACAGCGATGAAG 3161
Qy 2988 TAGCTTCATATTTGCTTCAATACCAGAACCAATTTCACTCTCTAGGTTTTTAAGTGGC 3047
Db 3162 ATCCATTTGTGCAATTTACGTTGATGTCGAAGTTATAAGCTTCGAGATCCACTAGTAAC 3221
Qy 3048 TCGTGGCGAATTCATCCCTCAGGATATAGTATGTTTTCGCTTTTTCATAGGAGGGGAAA 3107
Db 3222 GCGCGCAGTGTCTGGAATTCG-----CTTGTGACATCTAGGGGGCCNATTCGCC 3276
Qy 3108 TGTAGTCTTATGCAATACTCTTGTAGTCTTTCGAACATGTTGTAACGATGAGTTAGCAATG 3167
Db 3277 CCTCTCCCCCCCCCTAACGTTTACTGGCCGAAGCGCTTGGAAATAAGCCGGTGTGTG 3336
Qy 3168 CTTTACAGGAGAGAAACACGTCGATCGGATTTGGTGGGAATGAGTGGTGTACATC 3227
Db 3337 TTTGCTATATGTGATTTTCCACCATAATTCGCCGCTTTTGGCAA-----TGTAGGGGCC 3390

```
Qy 3228 GTGCTTATTAGGAAGCAACAGACGGCTGTGACATGATTTGGACGAACCACTGAATTCC 3287
Db
Qy 3391 CGGAACCTGGCCCTGTCTTCTTGACGAGCAATTCCTAGGGTCTTTCCCTCTCGCCAAA 3450
Db
Qy 3288 GCATTGACAGAGATATTGTATTTAAATGCTAGCTCGATACAGCAAAAGCCATTGACCAT 3347
Db
Qy 3451 GGAATGCAAGGTCTGTTGAATGTGTAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAAGA 3510
Db
Qy 3348 TCACCAATTTGGTGTGACCTTCAAGCTTCAAGCTGCGGCAAGCACTCAGGGCCGAAGGG 3407
Db
Qy 3511 CAACAACAGCTGTAGCGACCTTTGACGGCAGCGGAACCCCACTGCGGCAAGGTGC 3570
Db
Qy 3408 CTGCTAAAGGAAGCGGAACACAGTAGAAGCCAGTCCGACAGAAAGCGTGTGACCCCGAT 3467
Db
Qy 3571 CTCTGCGGCCAAAAGCCACGTTATAGATACACTCTGAAAGGGCGGCAACACCCAGTGC 3630
Db
Qy 3468 GAATGTGACGTACTGGGCTATCTGGACA - AGGGAAAACGCAAGCGCAAGAGAGAAGCAGG 3526
Db
Qy 3631 CACGTTGTGAGTTGGATAGTTGTGMAAGAGTCAATGGCTCTCTCAAGGTAGTCAAC 3690
Db
Qy 3527 TAGCTTGACGTGGCTTACATGGCGATAGCTAGCTGGGCGGTTTTATGACACGAAGCG 3586
Db
Qy 3691 AAGGGGCTGAAGGATGCCCAAGGTACCCCAATGTATGGAATCTGATCTGGGGCCCTCG 3750
Db
Qy 3587 AACCGGAATTGCCAGCTGGGGCCCTCTGTAAAGTTGGAGGCCCTGCAAAATAACT 3646
Db
Qy 3751 GTGCACATGCTTTACATGTGTTTATGTCAGGTTTAAAAAGCTTAGGCCCCCGCAACCAC 3810
Db
Qy 3647 GGATGGCTTTCTTCCCGCAAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGA 3706
Db
Qy 3811 GGGACGTGGTTTCTTTGMAAACAACGATGATAAGCTTGCCACAACCCGGGATATTC 3870
Db
Qy 3707 CAGGATGAGGATCGTTTCGATGATTAAGCAAGATGATGACGCAAGTTCTCCGCCG 3766
Db
Qy 3871 CTGCAGCAATATGGGATCGGCCATTAACAAGATGATGACGCAAGTTCTCCGCCG 3930
Db
Qy 3767 CTTGGGTGGAGAGCTATTGGCTATGACTGGGCACACAGACATCGGTGCTCTGATG 3826
Db
Qy 3931 CTTGGGTGGAGAGCTATTGGCTATGACTGGGCACACAGACATCGGTGCTCTGATG 3990
Db
Qy 3827 CCGCGCTGTTCCGGCTGTACGCGCAGGGGCCCGGTTCTTTTGTCAAGACCGACCTGT 3886
Db
Qy 3991 CCGCGTGTTCGGCTGTACGCGCAGGGGCCCGGTTCTTTTGTCAAGACCGACCTGT 4050
Db
Qy 3887 CCGGTGCCCTGAATGAATGTGACGACGAGGCGCGGCTATCGTGGCTGGCCACGACGG 3946
Db
Qy 4051 CCGGTGCCCTGAATGAATGTGACGACGAGGCGCGGCTATCGTGGCTGGCCACGACGG 4110
Db
Qy 3947 GCGTTCTTGGCAGCTGTGCTCGACCTTGTCACTGAAGCGGGAAGGACTGGCTGTAT 4006
Db
Qy 4111 GCGTTCTTGGCAGCTGTGCTCGACCTTGTCACTGAAGCGGGAAGGACTGGCTGTAT 4170
Db
Qy 4007 TGGCGGAAGTGC CGGGCAGGATCTCTGTCTCATCTCACTTGTCTTCTCGCGCAGAAAATAT 4066
Db
Qy 4171 TGGCGGAAGTGC CGGGCAGGATCTCTGTCTCATCTCACTTGTCTTCTCGCGCAGAAAATAT 4230
Db
Qy 4067 CCATCATGGCTGATGCAATGCGGCGGTGACATCGCTTGATTCGGGCTACCTGCGCCATTCG 4126
Db
Qy 4231 CCATCATGGCTGATGCAATGCGGCGGTGACATCGCTTGATTCGGGCTACCTGCGCCATTCG 4290
Db
Qy 4127 ACCACCAAGCGAATCATCGATCCAGCGACGACGTACTCGGATGGAAGCCGGTCTTGTG 4186
Db
Qy 4291 ACCACCAAGCGAATCATCGATCCAGCGACGACGTACTCGGATGGAAGCCGGTCTTGTG 4350
Db
Qy 4187 ATCAGGATGATCTGGAAGAGCATCAGGGGCTCGCGCCAGCGCAACTGTTCGCCAGGC 4246
Db
Qy 4351 ATCAGGATGATCTGGAAGAGCATCAGGGGCTCGCGCCAGCGCAACTGTTCGCCAGGC 4410
Db
Qy 4247 TCAGGCGCGCATGCCCGACGGCGAGGATCTCGTGTGACCCCAATGGCGATGCCCTGTTC 4306
Db
Qy 4411 TCAGGCGCGCATGCCCGACGGCGATGATCTCGTGTGACCCCAATGGCGATGCCCTGTTC 4470
```

```
Qy 4307 CGAATATCATGTTGGAAAATGGCGCTTTTCTGATTTCATGACTGTGGCGGCTGGTG 4366
Db
Qy 4471 CGAATATCATGTTGGAAAATGGCGCTTTTCTGATTTCATGACTGTGGCGGCTGGTG 4530
Db
Qy 4367 TGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCG 4426
Db
Qy 4531 TGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCG 4590
Db
Qy 4427 GCGAATGGGCTGACCGCTTCTCTGTCCTTTACGGTATCGGCTCCCGATTCGCGAGCGCA 4486
Db
Qy 4591 GCGAATGGGCTGACCGCTTCTCTGTCCTTTACGGTATCGGCTCCCGATTCGCGAGCGCA 4650
Db
Qy 4487 TCGCCCTTCTATCGCTTCTTTGACGAGTCTTCTGAGCGGACCTCTGGGTTCCGATAAAT 4546
Db
Qy 4651 TCGCCCTTCTATCGCTTCTTTGACGAGTCTTCTGAGCGGACCTCTGGGTTCCGATAAAT 4710
Db
Qy 4547 AAAAGATTTTATTAGTCTCCAGAAAAGGGGGGAATGAAAGACCCCACTGTAGGTTTG 4606
Db
Qy 4711 AAAAGATTTTATTAGTCTCCAGAAAAGGGGGGAATGAAAGACCCCACTGTAGGTTTG 4770
Db
Qy 4607 GCAAGCTAGCTTAAGTAAAGCCATTTTGCAGGCAATGGAAGGATGGAAGGATGGAAGTAG 4666
Db
Qy 4771 GCAAGCTAGCTTAAGTAAAGCCATTTTGCAGGCAATGGAAGGATGGAAGGATGGAAGTAG 4830
Db
Qy 4667 AGAAGTTTCAATCAAGCTCAGGAACAGATGGAAGGATGGAAGGATGGAAGGATGGAAGTAG 4726
Db
Qy 4831 AGAAGTTTCAATCAAGCTCAGGAACAGATGGAAGGATGGAAGGATGGAAGGATGGAAGTAG 4890
Db
Qy 4727 CTGTGGTAAAGCAGTTCCTGCCCCGGCTCAGGGCCAAAGAACAGATGGAAGGATGGAAGTAG 4786
Db
Qy 4891 CTGTGGTAAAGCAGTTCCTGCCCCGGCTCAGGGCCAAAGAACAGATGGAAGGATGGAAGTAG 4950
Db
Qy 4787 GGCACAAACAGGATATCTGTGGTAAAGCAGTTCCTGCCCCGGCTCAGGGCCAAAGAACAGATG 4846
Db
Qy 4951 GGCACAAACAGGATATCTGTGGTAAAGCAGTTCCTGCCCCGGCTCAGGGCCAAAGAACAGATG 5010
Db
Qy 4847 GTCCCCAGATGCGGTCAGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGT 4906
Db
Qy 5011 GTCCCCAGATGCGGTCAGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGT 5070
Db
Qy 4907 GCGCCAGGACCTGAAATGACCTGTGCTTATTTGAACCTAAACCAATCAGTTCGCTTCTC 4966
Db
Qy 5071 GCGCCAGGACCTGAAATGACCTGTGCTTATTTGAACCTAAACCAATCAGTTCGCTTCTC 5130
Db
Qy 4967 GCTTCTGTTCCGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTC 5026
Db
Qy 5131 GCTTCTGTTCCGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTC 5190
Db
Qy 5027 GGGGCGCCAGTCTCTCGATTGACTGAGTCGCGGGTACCCGTGTATCCAATAAACCTC 5086
Db
Qy 5191 GGGGCGCCAGTCTCTCGATTGACTGAGTCGCGGGTACCCGTGTATCCAATAAACCTC 5250
Db
Qy 5087 TTGCAGTTGCATCCGACTTGTGCTTCTGCTTCTTGGAGGGTCTCTCTGAGTGATT 5146
Db
Qy 5251 TTGCAGTTGCATCCGACTTGTGCTTCTGCTTCTTGGAGGGTCTCTCTGAGTGATT 5310
Db
Qy 5147 GACTACCGCTCAGCGGGGCTCTTTCATTGG 5177
Db
Qy 5311 GACTACCGCTCAGCGGGGCTCTTTCATTGG 5341
```

RESULT 6

```
Q879093 LOCUS 7436 bp DNA linear PAT 11-OCT-2004
DEFINITION Sequence 6 from Patent WO2004062554.
ACCESSION Q879093
VERSION Q879093.1 GI:54033470
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 other sequences; artificial sequences.
AUTHORS Martinez-Serrano,A., Liste,I. and Villa,A.
```


Query Match		38.1%; Score 1972.6; DB 6; Length 7436;		
Best Local Similarity		65.1%; Pred. No. 0;		
Matches 3371; Conservative		0; Mismatches 1529; Indels 281; Gaps 20;		
Qy	1	TTTGAAGACCCCA	CCCGTAGTGGCAAGCTAGCTTAAGTAAACGCCACTTTGCAAGCAT	60
Db	1	TTTGAAGACCCCA	CCCGTAGTGGCAAGCTAGCTTAAGTAAACGCCACTTTGCAAGCAT	60
Qy	61	GGAAATAATACAT	CTCAGAAATGAAAGTTTCAGATCAAGTCAAGGACAAAGAAACAGC	120
Db	61	GGAAATAATACAT	CTCAGAAATGAAAGTTTCAGATCAAGTCAAGGACAAAGAAACAGC	120
Qy	121	TGAATACCAACAGG	ATATCTGTGTTAAGCGGTTTCCTGCCCGGCTCAGGGCCCAAGAACA	180
Db	121	TGAATACCAACAGG	ATATCTGTGTTAAGCGGTTTCCTGCCCGGCTCAGGGCCCAAGAACA	180
Qy	181	GATGAGACAGCTG	ATGATGGGCAACAGGATATCTGTGTAAAGCAAGTATCTCTGCCCCCGG	240
Db	181	GATGAGACAGCTG	ATGATGGGCAACAGGATATCTGTGTAAAGCAAGTATCTCTGCCCCCGG	240
Qy	241	CTCGGGCCCAAGAA	CAGATGGTCCCAAGTCCAGCGCTCAGCGCTCAGCGATTTCTAGTGAA	300
Db	241	CTCGGGCCCAAGAA	CAGATGGTCCCAAGTCCAGCGCTCAGCGCTCAGCGATTTCTAGTGAA	300
Qy	301	TCATCAGATGTTTC	CAGGGTGCCCAAGGACCTGAAATGACCCCTGACCTTATTTCGAAC	360
Db	301	TCATCAGATGTTTC	CAGGGTGCCCAAGGACCTGAAATGACCCCTGACCTTATTTCGAAC	360
Qy	361	TAAACCAATCAGTT	CGCTTCTCGTTCCTGTTTCGGCGGCTTCGGCTCTCCGAGCTCAATAAA	420
Db	361	TAAACCAATCAGTT	CGCTTCTCGTTCCTGTTTCGGCGGCTTCGGCTCTCCGAGCTCAATAAA	420
Qy	421	AGAGCCACAACCC	CTCAGTCGGCGGCCAGTCTTCCGATAGACTGCGTCGCCCGGGTAC	480
Db	421	AGAGCCACAACCC	CTCAGTCGGCGGCCAGTCTTCCGATAGACTGCGTCGCCCGGGTAC	480
Qy	481	CCGTATTTCCAAATA	AGCCCTCTGCTGTTTTCATCCGAATCGTGGTCTCGCTGTTCCCTTG	540
Db	481	CCGTATTTCCAAATA	AGCCCTCTGCTGTTTTCATCCGAATCGTGGTCTCGCTGTTCCCTTG	540
Qy	541	GGAGGGTCTCCTCT	GAGTGAATGACTACCAACGACGGGGTCTTTTCATTTGGGGGCTCGT	600
Db	541	GGAGGGTCTCCTCT	GAGTGAATGACTACCAACGACGGGGTCTTTTCATTTGGGGGCTCGT	600
Qy	601	CCGGATTTGGAGAC	CCCTCCAGGACCAACCGACCCACACCGGAGAGTAAAGTGGCC	660
Db	601	CCGGATTTGGAGAC	CCCTCCAGGACCAACCGACCCACACCGGAGAGTAAAGTGGCC	660
Qy	661	AGCAACTTATCTGT	GTCTGTCAGTGTCTATGTTGATGTTATGCGCCTGGC	720
Db	661	AGCAACTTATCTGT	GTCTGTCAGTGTCTATGTTGATGTTATGCGCCTGGC	720
Qy	721	TCGTACTAGTATAG	CTAACTCTGTAATCTGCGGACCCGTGGTGGAACTGACGAGTT	780
Db	721	TCGTACTAGTATAG	CTAACTCTGTAATCTGCGGACCCGTGGTGGAACTGACGAGTT	780
Qy	781	CTGAACACCCGGCC	CAACCTCTGGAGACGTCCACAGGACTTTGGGGGGCGTTTTTGTGG	840
Db	781	CTGAACACCCGGCC	CAACCTCTGGAGACGTCCACAGGACTTTGGGGGGCGTTTTTGTGG	840
Qy	841	CCCGACTGAGGAAG	GGGAGTCGATGTGGAATCCGACCCCGTCAAGGATATGTGGTCTTGGT	900

Db	3864	GAACAATCGCATCGAGGAGCAGGTA	CTCGGATGGAAGCCGGTCTTGTCATCAGGATGA	3922	
Qy	4197	TCTGGAAGAAAGAGCATCAGGGGGT	TCGGCCAGGCGCAACTCTGTC	CCAGGCTCAAGGCGCG	4256
Db	3924	TCTGGAAGAAAGAGCATCAGGGGGT	TCGGCCAGGCGCAACTCTGTC	CCAGGCTCAAGGCGCG	3983
Qy	4257	CATGCCCGAGCGGAGGATCTCGT	GTGAGCCCATGCGGATATCAT	TGCGGAATATCAT	4316
Db	3984	CATGCCCGAGCGGAGGATCTCGT	GTGAGCCCATGCGGATATCAT	TGCGGAATATCAT	4043
Qy	4317	GGTGGAAAATGGCCGCTTTCTG	ATTCATCGACTGTGGCCGGCTGGGT	GTGCGGACCG	4376
Db	4044	GGTGGAAAATGGCCGCTTTCTG	ATTCATCGACTGTGGCCGGCTGGGT	GTGCGGACCG	4103
Qy	4377	CTATCAGGACATAGCGTTGGCT	ACCCCGTGATTTGCTGAAAGACT	TTCGGCGCGCAATGGGC	4436
Db	4104	CTATCAGGACATAGCGTTGGCT	ACCCCGTGATTTGCTGAAAGACT	TTCGGCGCGCAATGGGC	4163
Qy	4437	TGACCGCTTCTCGTGTCTTTAC	GGATATCGCGCTTCGCGATTCG	GAGCGCATGCCCTTCTA	4496
Db	4164	TGACCGCTTCTCGTGTCTTTAC	GGATATCGCGCTTCGCGATTCG	GAGCGCATGCCCTTCTA	4223
Qy	4497	TCGCCTCTTTGACGAGTCTTCT	TGACGGGACCTCTGGGGTTC	GATATAAATAAAGATTTT	4556
Db	4224	TCGCCTCTTTGACGAGTCTTCT	TGACGGGACCTCTGGGGTTC	GATATAAATAAAGATTTT	4283
Qy	4557	ATTTAGTCTCCAGAAAAGGGG	GAATGAAGACCCCACTGT	PAGGTTTGGCAAGCTAGC	4616
Db	4284	ATTTAGTCTCCAGAAAAGGGG	GAATGAAGACCCCACTGT	PAGGTTTGGCAAGCTAGC	4343
Qy	4617	TTAAGTAAACCCCATTTTTC	GAAGGATGGAATAATACAT	CTGAGAAATAGAGAATTCAG	4676
Db	4344	TTAAGTAAACCCCATTTTTC	GAAGGATGGAATAATACAT	CTGAGAAATAGAGAATTCAG	4403
Qy	4677	ATCAAGTCAAGAACAGATGGA	ACAGTGAATATGGCCAAAC	CAGGATATCTGTGCTAG	4736
Db	4404	ATCAAGTCAAGAACAGATGGA	ACAGTGAATATGGCCAAAC	CAGGATATCTGTGCTAG	4463
Qy	4737	CAGTTCCTGCCCCGGCTCAG	GGCCAAAGACAGATGGAAC	CAGCTGAATATGGGCCAAACAG	4796
Db	4464	CAGTTCCTGCCCCGGCTCAG	GGCCAAAGACAGATGGAAC	CAGCTGAATATGGGCCAAACAG	4523
Qy	4797	GATATCTGTGGTAAGCAGTTC	TGCCCCGGCTCAGGGCCAA	GAAGACAGATGGTCCCCAGAT	4856
Db	4524	GATATCTGTGGTAAGCAGTTC	TGCCCCGGCTCAGGGCCAA	GAAGACAGATGGTCCCCAGAT	4583
Qy	4857	GCGGTCAGCCCTCAGCAGTTC	TCAGAGACCAATCAGATG	TTTCCAGGGTGCCTCCAGGA	4916
Db	4584	GCGGTCAGCCCTCAGCAGTTC	TCAGAGACCAATCAGATG	TTTCCAGGGTGCCTCCAGGA	4643
Qy	4917	CCTGAAATGACCTGTGCTTT	ATTGAACTAAACCAATCAG	TTCCGCTTCTCGCTTCTGTTTC	4976
Db	4644	CCTGAAATGACCTGTGCTTT	ATTGAACTAAACCAATCAG	TTCCGCTTCTCGCTTCTGTTTC	4703
Qy	4977	GCGCGCTTCTGCTCCCGAG	CTCAATAAAGACCCCAAC	CCCTCACTCGGGGCGCCAG	5036
Db	4704	GCGCGCTTCTGCTCCCGAG	CTCAATAAAGACCCCAAC	CCCTCACTCGGGGCGCCAG	4763
Qy	5037	TCCTCCGATTTGACTGAGT	CGCCGGGTACCCGGTGAT	CCATAAACCCTCTTGCAGTTCG	5096
Db	4764	TCCTCCGATTTGACTGAGT	CGCCGGGTACCCGGTGAT	CCATAAACCCTCTTGCAGTTCG	4823
Qy	5097	ATCCGACTTGTGCTCGCTGT	TCTTTGGGAGGGTCTCCT	CTGAGTGAATGACTACCCGT	5156
Db	4824	ATCCGACTTGTGCTCGCTGT	TCTTTGGGAGGGTCTCCT	CTGAGTGAATGACTACCCGT	4883
Qy	5157	CAGCGGGGGTCTTTTCATTT	GG	5177	
Db	4884	CAGCGGGGGTCTTTTCATTT	GG	4904	

RESULT 7

Db 2685 TTCCCTGGATGTTTTCGGACCAAGTGGTGAATTCAGTGGATCCACTAGTAAAGCGCCCA 2744
Qy 2699 AGGTACGGGTCCAAATATATGGCTGAATGAAGTATTTTGTTCGGAAAGAGTCAATCCATGA 2758
Db 2745 GTGTGCTGGAATTAATTCGCTGTCTGCGAGGCGCGCTGTTGGGTGAGTACTCCCTCTC 2804
Qy 2759 AGAGTGCAGAATTAGACAGTGGGGTGTGAGAGCTGTTCGCACGACGAGAGTGTGGGG 2818
Db 2805 AAA-----AGCGGGCATGACTTTCGCGCTAAGAT-----TG 2835
Qy 2819 TCACCTTTCGACCTACATAAATCATCAATTTTCAATTCACATTTTAACTGTTATAAAG 2878
Db 2836 TCAGTTTCCAAAACGAGGAGATTGATATCA-----CCTGGCCCGCGGTGATG 2886
Qy 2879 TGATTTTTTCTTGTCTTCACTAAATTCAGCTTAATTAATTAATTAAGAACTAAGAAAT 2938
Db 2887 CCTTTGAGGTGGCGCGTCCATCTGGTTCAGAAAGACAACTCTTTTGTGTCAAGCTTG 2946
Qy 2939 TTATCCACAGAAAGGAATATTAATAATCACTGATTAACATATAAATAGTCTCATAT 2998
Db 2947 AGGTGTGCGAGCTTGAGATCTGGCCATACACTTGAAGTGAC-----AATGACATCCACT 3000
Qy 2999 TTGCTTCAAAATACAGAACCAATTTCAACTTCTCTAGGTTTTTAAGTGGCTCGTCCGGAAT 3058
Db 3001 TTGCTTCTCTCCACAGGTGTCACTCCAGGTCCAACTCCAGGTGATCGACATGCA 3060
Qy 3059 TGATCCCTCAGGATATAGTATGTTTTCGCTTTTGCATAGGAGGGGGAATAGTCTTTAT 3118
Db 3061 TCTAGGGGGGCAATTCGCCCCCTCTCCCTCCCCCCCCCTAAAGTTACTGGCCGAGCGG 3120
Qy 3119 GCAATACTCTTGTAGTCTTGCAACATGATACGATAGTATAGCAACATGCTTCAAGGA 3178
Db 3121 CTTGGAAATAGCGCGGTGTGTGTGTCTATATGTATGTATTTCCACCAT----- 3168
Qy 3179 GAGAAAAGCACCGTGCAATGCGGATTTGTTGAAGTAAAGTGGTACGATCGTCTTATTA 3238
Db 3169 -----ATTGCGCTCTTTTGGCAATGTGAGGCGCCGGAACCTGCGCCCTGTCTTCTTG 3220
Qy 3239 GGAAGGCAACAGACGGGTCTGACATGATTTGAGCAACCACTGAATTCGCAATTCGACAG 3298
Db 3221 ACGAGCATTCCTAGGGGTCTTTCCCTCTCGCAAGGAATGCAAGGTCTGTGTAATGTC 3280
Qy 3299 ATATTGTATTAAGTCTAGCTCGATACAGCAAAAGCGCATTTGACCAATTCACCAATTG 3358
Db 3281 GTGAAGGAAGCAGTTCTCTGGAAGCTTCTTGAAGACAAACAAAGCTCTGTAGCACCCCT 3340
Qy 3359 GTGTGCACCTCCAGCTTTCAGCTGCGCAGCACTCAGGGCGCAAGGGCTGTAAAGGA 3418
Db 3341 TGCAGGCAAGGGAACCCCACTGCGGCAAGGTGCTCTCGCGCCAAAGCCACGTGTA 3400
Qy 3419 AGCGGAACACGTAGAAAGCCAGTCCGCAAGAACGGTGTGACCCCGGATGAATGTACGCT 3478
Db 3401 TAAGATACACCTGCNAAGCGGCAC-----AACCCAGTGCCACAGTTGTGAGTTGGATAG 3455
Qy 3479 ACTGGCTATCTGGAAGGGAAGAAAGCAAGCGCAAGAGAGAAAGAGGTAGTTCAGATG 3538
Db 3456 TTGTGGAAGAGTCAATAGGCTCTCTCAAGCG-----TAGTCAACAGGGGCTGAAGGAT 3511
Qy 3539 GGCTTACATGCGGATAGTACACTGGGCGGTTTTATGACAGACAGCAAGCAAGCAAGTTCG 3598
Db 3512 GCCCAGAAAGTACCCCAATTTATGGGAATCTGATCTGGGGCTCGGTGCACATGCTTTAC 3571
Qy 3599 CAGCTGGGGGCCCTCTGTAAGGTGGGAAGCCCTGCAAGTAACTGAGTGGCTTCT 3658
Db 3572 ATGTGTTTAGTCGAGGTTAAAGAGCTTAGGCCCCCGGAA-----CCACGGGACGTGGTTT 3629
Qy 3659 TGCGCCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGAGGAT 3718
Db 3630 TCCTTTGAAGAAACACGATGATAGCTTCCACAAACCCCGGGATTAATCTTCGACCCAATA 3689
Qy 3719 CGTTTCGATGATTGAACAAGATGATTTGACGCAAGGTTCTCCGCGCCGCTTGGGTGGAGA 3778
Db 3690 TGGGATCGGCCAATTGAAACAAGATGATTTGACGAGGTTCTCCGCGCCGCTTGGGTGGAGA 3749

Qy 3779 GGCTATTTTCGGCTATGACTGGGCAAAACAGACAATTCGGCTGCTCTGATGCGCCCGTGTTC 3838
Db 3750 GGCTATTTTCGGCTATGACTGGGCAAAACAGACAATTCGGCTGCTCTGATGCGCCCGTGTTC 3809
Qy 3839 GGCTGTCAAGCGCAGGGGCGCCGGTCTTTTGTCAAGACCGACCTGTTCGGTGCCTTGA 3898
Db 3810 GGCTGTCAAGCGCAGGGGCGCCGGTCTTTTGTCAAGACCGACCTGTTCGGTGCCTTGA 3869
Qy 3899 ATGAACTGCAAGCGCAGGGCAGCGGCTATCTGTGGCTGGCCACAGCGGGCTTCTTTGCG 3958
Db 3870 ATGAACTGCAAGCGCAGGGCAGCGGCTATCTGTGGCTGGCCACAGCGGGCTTCTTTGCG 3929
Qy 3959 CAGCTGTCTCGACGTTGTCACTGAAGCGGAAAGGACTGCTGCTATTTGGGCGCAAGTGC 4018
Db 3930 CAGCTGTCTCGACGTTGTCACTGAAGCGGAAAGGACTGCTGCTATTTGGGCGCAAGTGC 3989
Qy 4019 CGGGCAGGATCTCTCTGCTATCTCACCTGCTCTGCGGAGAAAGTATCCATCATGSGCTG 4078
Db 3990 CGGGCAGGATCTCTCTGCTATCTCACCTGCTCTGCGGAGAAAGTATCCATCATGSGCTG 4049
Qy 4079 ATGCAATTCGCGCGCTGCATACGCTTCGATCCGGTACCTGCCATTCGCCACCAAGCGA 4138
Db 4050 ATGCAATTCGCGCGCTGCATACGCTTCGATCCGGTACCTGCCATTCGCCACCAAGCGA 4109
Qy 4139 AACATCGCATCGACGAGCAGTACTCGGATGGAAGCCGCTTTGTGATCAGATGATC 4198
Db 4110 AACATCGCATCGACGAGCAGTACTCGGATGGAAGCCGCTTTGTGATCAGATGATC 4169
Qy 4199 TGAACGAAGACATCAGGGCTCGCGCAGCGCAACTGTTGCGCAGGCTCAAGGCGGCA 4258
Db 4170 TGAACGAAGACATCAGGGCTCGCGCAGCGCAACTGTTGCGCAGGCTCAAGGCGGCA 4229
Qy 4259 TGCCCGACGCGAGGATCTCGTCTGTAGCCATGCGGATGCTGCTGCGCAATATCATGG 4318
Db 4230 TGCCCGACGCGAGGATCTCGTCTGTAGCCATGCGGATGCTGCTGCGCAATATCATGG 4289
Qy 4319 TGGAAAATGCGCGCTTTTGTGATTCGACTGTGCGCGCTGGGTGTGCGGACCGCT 4378
Db 4290 TGGAAAATGCGCGCTTTTGTGATTCGACTGTGCGCGCTGGGTGTGCGGACCGCT 4349
Qy 4379 ATCAGGACATAGCGTGGCTACCGCTGATTTGCTGAAGAGCTTGGCGGCGAATGGGTG 4438
Db 4350 ATCAGGACATAGCGTGGCTACCGCTGATTTGCTGAAGAGCTTGGCGGCGAATGGGTG 4409
Qy 4439 ACCGTTCTCTGCTTTTACGCTATCGCGCTCCCGATTCGACGCGATCGCTTCTATC 4498
Db 4410 ACCGTTCTCTGCTTTTACGCTATCGCGCTCCCGATTCGACGCGATCGCTTCTATC 4469
Qy 4499 GCCTTCTTGAACGATTC-----TTC 4518
Db 4470 GCCTTCTTGAACGATTCCTGCTGAGGCGGATCTGATCAAGACAGGATGAGGATCGTTT 4529
Qy 4519 TGAGCGGACCTCTGGGTTCGATAAATAAAGATTTTATTTAGTCTCCAGAAAAGGG 4578
Db 4530 CGCGCGGACTCTGGGTTCGATAAATAAAGATTTTATTTAGTCTCCAGAAAAGGG 4589
Qy 4579 GGAATGAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAAAGCCATTTTGAAG 4638
Db 4590 GGAATGAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAAAGCCATTTTGAAG 4649
Qy 4639 GCATGGAATAATACATAACTGAGAATAGAGAATTCAGATCAAGGTCAAGAACAGATGGA 4698
Db 4650 GCATGGAATAATACATAACTGAGAATAGAGAATTCAGATCAAGGTCAAGAACAGATGGA 4709
Qy 4699 ACAGCTGAATATGAGGCGCAACAGGATATCTGTGTAAGCAGTTCCTGCGCCCGCTCAGGG 4758
Db 4710 ACAGCTGAATATGAGGCGCAACAGGATATCTGTGTAAGCAGTTCCTGCGCCCGCTCAGGG 4769
Qy 4759 CCAGACAGATGGAACAGCTGATATGGGCCAAACAGGATATCTGTGTAAGCAGTTCCT 4818
Db 4770 CCAGACAGATGGAACAGCTGATATGGGCCAAACAGGATATCTGTGTAAGCAGTTCCT 4829

Qy 4819 TGCCCCGGCTCAGGGCCCAAGAACAGATGTTCCCGAGATGCGGTCCAGCCCTCAGCAGTTT 4878
Db 4830 TGCCCCGGCTCAGGGCCCAAGAACAGATGTTCCCGAGATGCGGTCCAGCCCTCAGCAGTTT 4899
Qy 4879 CTAGAGAACCATCAGATGTTCCAGGGTGCCCCAAGAGACCTGAAATGACCTGTGCGCTTA 4938
Db 4890 CTAGAGAACCATCAGATGTTCCAGGGTGCCCCAAGAGACCTGAAATGACCTGTGCGCTTA 4949
Qy 4939 TTTGGAATCAACCATCAGTTCGCTTCTCGTTCCTGTTGCGCGGCTTCTGCTCCCGGAGCT 4998
Db 4950 TTTGGAATCAACCATCAGTTCGCTTCTCGTTCCTGTTGCGCGGCTTCTGCTCCCGGAGCT 5009
Qy 4999 CAATAAAGAGCCCAACACCCCTCACTCGGGCGCCAGTCCCTCGAATGACTGAGTCCGCC 5058
Db 5010 CAATAAAGAGCCCAACACCCCTCACTCGGGCGCCAGTCCCTCGAATGACTGAGTCCGCC 5069
Qy 5059 CGGGTACCCGTGATCCAAATAAACCCCTTTCGAGTTGCATCCGACTTGTGTGCTCGCTGT 5118
Db 5070 CGGGTACCCGTGATCCAAATAAACCCCTTTCGAGTTGCATCCGACTTGTGTGCTCGCTGT 5129
Qy 5119 TCCTTGGAGGGTCTCCTCTGAGTGAATGACTACCCGTCAGCGGGGGTCTTTCATTTGG 5177
Db 5130 TCCTTGGAGGGTCTCCTCTGAGTGAATGACTACCCGTCAGCGGGGGTCTTTCATTTGG 5188

RESULT 8

BD138564 7165 bp DNA linear PAT 18-SEP-2002
LOCUS Expression system of nucleotide with reduced immunogenicity for use
in gene therapy.

ACCESSION BD138564

VERSION BD138564.1 GI:23233509

KEYWORDS JP 2002508976-A/3.

SOURCE synthetic construct

ORGANISM other sequences; artificial sequences.

REFERENCE 1 (bases 1 to 7165)

AUTHORS Radosevich, T.J. and Jr, C.J.L.

TITLE Expression system of nucleotide with reduced immunogenicity for use

JOURNAL in gene therapy

COMMENT Patent: JP 2002508976-A 3 26-MAR-2002;

HUMAN GENE THERAPY RESEARCH INSTITUTE

OS Artificial Sequence

PN JP 2002508976-A/3

PD 26-MAR-2002

PF 13-JAN-1999 JP 2000540263

PR 14-JAN-1998 US 60/071409

PI THOMAS J RADOSEVICH, CHARLES J LINK JR

PC C12N15/09,A61K31/711,A61K48/00,A61P37/06,C12N1/15,C12N1/19, PC

C12N1/21.

PC C12N5/10//A61K35/76,(C12N15/09,C12N1/93),C12N15/00,C12N5/00,

PC (C12N15/00,C12N1/93)

CC Description of Artificial Sequence : recombinant vector FH

Key Location/Qualifiers

FT source 1..7165

FT Location/Qualifiers

1..7165 /organism='Artificial Sequence'.

/organism='synthetic construct'

/mol_type='genomic DNA'

/db_xref='taxon:32630'

FEATURES

source

ORIGIN

Query Match 31.8%; Score 1644.6; DB 6; Length 7165;
Best Local Similarity 62.9%; Pred. No. 0;
Matches 3284; Conservative 0; Mismatches 1309; Indels 624; Gaps 23;

Qy 1 TTTGAAGACCCACCGCTAGTGCGAGCTAGCTTAAGTAACGCCACTTTGCAAGGCAT 60
Db 175 TTTGAAGACCCACCGCTAGTGCGAGCTAGCTTAAGTAACGCCACTTTGCAAGGCAT 234
Qy 61 GGAATAATACATAACTGAGAAATAGAAAAGTTTCAGATCAAGGTCAGGAAACAAAGAACAGC 120

Db 235 GGAATAATACATAACTGAGAAATAGAAAAGTTTCAGATCAAGGTCAGGAAACAAAGAACAGC 294
Qy 121 TGAATACCAACACAGGATATCTGTGTAAGCGGTTCCCTGCCCCGCTCAGGGCCCAAGACACA 180
Db 295 TGAATACCAACACAGGATATCTGTGTAAGCGGTTCCCTGCCCCGCTCAGGGCCCAAGACACA 354
Qy 181 GATGAGACAGCTGAGTGGGGCCAAACAGGATATCTGTGTAAGCGGTTCCCTGCCCCGG 240
Db 355 GATGAGACAGCTGAGTGGGGCCAAACAGGATATCTGTGTAAGCGGTTCCCTGCCCCGG 414
Qy 241 CTGGGGCCCAAGAACAGATGTTCCCGAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGA 300
Db 415 CTGGGGCCCAAGAACAGATGTTCCCGAGATGCGGTCCAGCCCTCAGCAGTTTCTAGTGA 474
Qy 301 TCATCAGATGTTCCAGGGTGCCCCCAAGGACCTGAAAATGACCCCTGACCTTATTTGAA 360
Db 475 TCATCAGATGTTCCAGGGTGCCCCCAAGGACCTGAAAATGACCCCTGACCTTATTTGAA 534
Qy 361 TAACCAATCAGTTTCGCTTCTCGTTCCTGTTGCGCGCTTCCGCTCTCCGAGCTCAATAA 420
Db 535 TAACCAATCAGTTTCGCTTCTCGTTCCTGTTGCGCGCTTCCGCTCTCCGAGCTCAATAA 594
Qy 421 AGAGCCCAACCCCTCACTCGGGCGCCAGTCTTCGAGTAGACTGCGTCCCGGGTAC 480
Db 595 AGAGCCCAACCCCTCACTCGGGCGCCAGTCTTCGAGTAGACTGCGTCCCGGGTAC 654
Qy 481 CCGTATTTCCCAATTAAGACCTTTCGCTGTTTGCATCCGAATCGTGGTCTCGCTGTTCTTG 540
Db 655 CCGTATTTCCCAATTAAGACCTTTCGCTGTTTGCATCCGAATCGTGGTCTCGCTGTTCTTG 714
Qy 541 GGAGGGTCTCCTCTGAGTGAATGACTACCAACAGGGGGTCTTTTCATTTGGGGGCTCGT 600
Db 715 GGAGGGTCTCCTCTGAGTGAATGACTACCAACAGGGGGTCTTTTCATTTGGGGGCTCGT 774
Qy 601 CCGGGATTTGAGACCCCTCCCGAGGAGCAACCGAACCAACCGAGGAGTAACTGGCC 660
Db 775 CCGGGATTTGAGACCCCTCCCGAGGAGCAACCGAACCAACCGAGGAGTAACTGGCC 834
Qy 661 AGCAACTTATCTGTGCTGTCGATGCTAGTGTCTATGTTGATCTGATGCGCCCTGCG 720
Db 835 AGCAACTTATCTGTGCTGTCGATGCTAGTGTCTATGTTGATCTGATGCGCCCTGCG 894
Qy 721 TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGCGGACCCGCTGGTGGAACTGACGAGTT 780
Db 895 TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGCGGACCCGCTGGTGGAACTGACGAGTT 954
Qy 781 CTGAACACCCCGCGCAACCCCTCGGAGACGTCCTCAGGACCTTTGGGGGCGCTTTTGTGG 840
Db 955 CTGAACACCCCGCGCAACCCCTCGGAGACGTCCTCAGGACCTTTGGGGGCGCTTTTGTGG 1014
Qy 841 CCGGACCTGAGGAGGAGGAGTGGATGGAATCCGACCCGCTCAGGATATGTTGTTCTGCT 900
Db 1015 CCGGACCTGAGGAGGAGGAGTGGATGGAATCCGACCCGCTCAGGATATGTTGTTCTGCT 1074
Qy 901 AGGAGACGAGAACTTAAACAGTTCCCGCTCTGAAATTTTGTCTTTCGGTTTGGAA 960
Db 1075 AGGAGACGAGAACTTAAACAGTTCCCGCTCTGAAATTTTGTCTTTCGGTTTGGAA 1134
Qy 961 CCGAAGCCCGCGCTTGTCTGTCGAGCAACGCTTGGGCTGCGAGTCTCTAGAGGA 1020
Db 1135 CCGAAGCCCGCGCTTGTCTGTCGAG----- 1162
Qy 1021 TCAATTCGGCACGAGTAAATCGGTGCTGCGCTTTTAGGACATATGAGTATGCGACAGT 1080
Db 1163 -----CGCTGACAGATCGTTCGTGTGTGTCTGTCTGTCTGCTGACTGTGT 1202
Qy 1081 GGGATGACTTTCTCTGATCAGAAAGAGGACACTGACAGCTGTACAGAGTCTGTGAAGTTG 1140
Db 1203 TTTGTGATTTGTCTGAAATTAAGGCCAGACTGTACC-----ACTCCCTTAAGTTG 1255
Qy 1141 ATGCTCGCTCAGTGACAGCTTTGCTTCTCCCTCCCATCTTAAATAATGGCCCAACTCTCAAG 1200
Db 1256 ACCTTAGGTCACCTGGAAGATGTGAGCGGATCGCTCACAACCAAGTCCGTAGATGTCAAG 1315

Qy	1201	AGAGATGAAGTCTTTATAAAACTGCACCTGATCACCCCTTTATCTCATTTGTGTGTTGTAGTTC	1260
Db	1316	AAGAGACGTTG-----GGTTACCTTCTGCTCTGCGAATGG	1351
Qy	1261	TCGTGCCCATCATTTGGCATAGTGCAGCTCAGCTTCCTGAAATGGGAACGAAGAATTGCA	1320
Db	1352	CCAACCTTTAAAGTCGAGTGCCTGGAGACGGCACCTTTAAACGAGACCTCATCA-----	1406
Qy	1321	CGGTTGGCTCAGTTTAATTCGACATATATCTCCAACTCCGGAAGCAAGSAAATGGCAGTG	1380
Db	1407	-----CCCAGGT	1413
Qy	1381	AAGATGAATGCAGATTTTCGAGAAGCTGTGATGGAACCGATGAGCAACATGGAAAGCAGAA	1440
Db	1414	TAAATCAAGGTCTTTTCACTGCGCCCGCATGGACACCCAGACAGGTCCTCTACATCGT	1473
Qy	1441	TCCAGTATCTTTTCAGATAATGAAGCCAAATCTCTCTAGATGCTTAAGAAATTTCCAAAAATTTCA	1500
Db	1474	GACCTGGGAGGCTTGGCTTTTGACCCCCCTCTCTGGGTCAAGCCCTTTGTACACCCATA	1533
Qy	1501	GCATAACAATGATCAAGAATTTAATGATGTTCTTTTCCAGCTAAATTCCTTACTTTCCCT	1560
Db	1534	GCCTCCGCCCTCTCTCTCCATCCGCCCGCTCTCTCCCTTGAACCTCTCTGTTCCGAC	1593
Qy	1561	CCATCCAGGACATGAGAAATCATAGGGGATATCTCCAAGTCATTTAGTAGGTCTGNAACA	1620
Db	1594	CCGCCCTCGATCTCTCTTTAT-----CCAGCCCTCACTCCTTCTCTTAGG	1638
Qy	1621	CCACAGTACTTTGATTTTCAGTTTCAGTATTTGAAACACTGAATGGCAGAGTCCAAGAGAATG	1680
Db	1639	CGCCGAAATTCGT-----NACT	1656
Qy	1681	CATTTAAACAACAGAGAGAGATGCGTAAATTTAGAGGAGCGTATATACAATGCAATCAGCAG	1740
Db	1657	CGAGATCCGGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAGTCCCCAGGCTCCCCAG	1716
Qy	1741	AAATTAAGTCTCTAGATGAAACCAAGTATATTTTGGACAGGAATAAAGGGGAATGA	1800
Db	1717	-----CAGGCAGAAGTATGCAAGCATGCATCTCAATTATGTACGCA	1757
Qy	1801	AACTGTTGAATAATATCACTAATGATCTGAGGCTGGAAGGATTTGGGAACAATCTCAGACAT	1860
Db	1758	ACCAGGTGTGGAAGTCCCCCAGGCTCCACAGGCGAGAGTAGTGCAA---GCATGCAT	1814
Qy	1861	TGAAAAATATCATCTTTACTCAAGGTGCCAGAAAGTCTCTCGTGACTGGGGAATTGAACCA	1920
Db	1815	CTCAATTAGTCAGCAACCATAGTCCCGCCCTAACTCCGCCCATCCCGCCCTAACTCCG	1874
Qy	1921	ACGATCTGGGCTCCAACATGACCATCGGGCTGTGACACAGCAGGTTGAATTCACAGGCA	1980
Db	1875	CCAGTTCCGCCCATCTCGCCGCCATGGCTGACTAATTTTTTTTATTTATGACAGAGGCC	1934
Qy	1981	CCTACATCAGCCGCTAACGCCACATCAAAATGAGATCAAAAGACTCACCACTGCATGGGA	2040
Db	1935	GAGSCCGCTCGGCTCTGAGCTATTCAGAGTAGTAGGAGGCT-----TTTTTGG	1987
Qy	2041	CACAAAACCATCAACAAGAGGACCCAGCCCATCTTTGGCTTACCGTCAATTTGGAAGT	2100
Db	1988	AGGCCTTAGGCTTTTTCAAAAAGCTTGGGCTGCAAGCTTGGTACCGAGCTCGGATCGATAT	2047
Qy	2101	TTTTCAGAGTCCACCATGTCTTTCAGGGCCAGATGCTTTCATAGACAGGAATGGGAGGAGG	2160
Db	2048	CTGGCGCGGTGACACGATGTCTGTGGGCCCTGG-----AAATGCGGACACCT	2096
Qy	2161	TCCTGAAGACCATGTGGCTGTCGGGTCAAGTGTTAATGATTTGGTGTGATGACTGGGAAG	2220
Db	2097	TCCTGGACCATATGCGGCTTTGGGCCCAGGACGTCAGCGCAGTACGCGATGA-----	2148
Qy	2221	CTACCGGCTGGCATCAACATCTTCACTCGCTCGCCACACAGAGAGGTAGTGAAGTG	2280
Db	2149	-----GATCAATAAAGGGGGGTGAGGA	2172

Qy	2281	ACCAAGTCTCTCTGGACTCCAGGTGAAAGAGAGATAGAGGCCCTCTCTGGACAAATATGG	2344
Db	2173	CCGGAGCGCGCCAGAAACCGCCGTGCACGACCCGCGAGCTCCCTCTGCTGCGCTCTCCCGG	2232
Qy	2341	TATACCAAGGCTTTCCAGGTCTAATAGTACTCCAGGTCTTTAAAGGTGATCGGGGGATCT	2400
Db	2233	GCTGCTGCCGAAATCGCCCCAACGACATCTTTGGGTGTGGCACATCGAAGAACCGCGCG	2292
Qy	2401	CTGTTTATCTGGAGTTCTGAGGATTCAGGACCAATCCAGGGAAGACGGGAACCCAGGAC	2460
Db	2293	GACCGTGACCGACAGTCCCGTAATCCGGTAACCCCGTTGAAATTCAGTGGATCC-----	2346
Qy	2461	TTAATGCAAAAGAGCCAGAACGGGAAAGAGGAGTGGAGCATGCAAGACATCTTA	2520
Db	2347	-----ACTA	2350
Qy	2521	ATACAGTCCGACTGGTGGGTGGCAGCGGCCCTTCACGAAGSCAGAGTGGAGATTTTTTCACG	2580
Db	2351	GTTACGGCCCGCAGTGTGCTGG-----	2372
Qy	2581	AAGSCCAGTGGGGTACGGTGTGTGACGACCGCTGGAACTGCGTGGAGGACTGGTGGTCT	2640
Db	2373	-----AATTAATTCGCTGTCTCGAGGGCGCGCTGTTGGGTGAGTACTCCCTCTCAAAA	2427
Qy	2641	GCAGGACTTGGGATACAAAGGTGTTCAAAGTGTGCATAAGCGAGCTTATTTTGGGAAAG	2700
Db	2428	CGCGGCATGACTTCTGCGCTAAGATTTGTCAAGTTTCCAAAACACGAGGAGATTG-----	2481
Qy	2701	GTACGGGTCCAATATGSGTGAATGAAGTATTTGTTTCGGGAAGAGTCAATCCATTGAAG	2760
Db	2482	-----ATATTCACTTGGCCCGC	2498
Qy	2761	AGTGCAGAAATAGACAGTGGGGTGTGAGAGCCTGTTTCGCACGACGAGAGATGCTGGGGGTC	2820
Db	2499	GGTGATGCTTTGAGGGTGGCGCGTCCATCTCGTCAGAAAAGACAACTCTTTTGTGTGTC	2558
Qy	2821	ACTTTCACCTCAATAAATGATCATATTTTTCATTCACATTTTTTAACTGTGTATAAAGTG	2880
Db	2559	AAGCTTGAGGTGTGCAGGCTTGAGATCTGGCCATACACTTGAAGTGAACAATGACATCCAC	2618
Qy	2881	ATTTTTTTCCTTTGCTTCACTAAATACAGCTTAATTAATTTTAAGNAACTAAGAAATTTT	2940
Db	2619	TTTGGCTTTCTCTCC-----ACAGGT	2639
Qy	2941	ATCCACAGAAAAGGAATATTTAAAAATCACTGGATAAACATATAAAATAGCTTTCATATTT	3000
Db	2640	GTCCACTCCAGTCCCACTGCAAGTGCATGAGCATGATCTAGGGCGCCAAATTCGCC	2699
Qy	3001	GCTTCAATAACCAAGAACCATTTCAACTTCTTAGTGTTTTAAAGTGCTCGTGCAGAAATG	3060
Db	2700	CCTCTCCCTCCCGCCCGCTTAAC-----GTTACTGGCCGA	2734
Qy	3061	ATCCCTCAGGATATAGTAGTTTCGCTTTTTCATAGGAGGGGGAATGTAGTCTTATGCT	3120
Db	2735	AGCGCTTGAATAAGCGCGGTGTGTGTTGCT-----ATATGTGATTTTCCAC	2784
Qy	3121	AATACTCTTGTAGTCTTGCAACATGTTAACAATGAGTTAGCAACATGCGCTTCAAGGAGA	3180
Db	2785	CATATTGCCGTCTTTTGGCAATGTAGGGCCCGGAACCTGGCCCTGTCTTCTTGACGAG	2844
Qy	3181	GA AAAAGCACCGTGCAATGCCGATTTGGTGGAAATGAGTGGTACGATCGTGCCCTTATTAGG	3240
Db	2845	CATTCTAGGGGTCTTTCCCTCTCGCCAAAGGA-----TGC	2882
Qy	3241	AAGCAACACAGCGGTCTGACATGGAATGGAAGCAACCACTGAATTCGGCATTCGACAGAT	3300
Db	2883	AAGGTCTGTTGAATGTGTGAAGGAAGCAGTCTCTCTGGAAGCTCTTGAAGACAAACAA	2942
Qy	3301	ATTGTAATTAAGTGCCTAGCTCGATACAGCAACGCCATTTGACCAATTCACCACTTGGT	3360
Db	2943	CGTCTGTAGCGACCCCTTTGACGACGCGGNAACCCCCCAACCTGGCGACAGGTGCTCTGGC	3002
Qy	3361	GTGCACCTCCAAAGCTTCAAGCTGCTCCGCAAGCACTCAGGGCGCAAGGGCTGCTAAAGGAAG	3420

Db 3003 GCCAAAGCCACGTGTATAGATACACCTGCAAGGGCGGCACA----- 3045
Qy 3421 CGGAACAGTGAAGCCAGTCCGAGAAACGCTGTCTGACCCCGGATGAATGTGAGTAC 3480
Db 3046 ---ACCCAGTGCACAGTGTGAGTTGGATAGTTTGGAAAGAGTCAATGGCTCTCCTC 3102
Qy 3481 TGGGCTATCTGGACAAGGGAAACGCAAGCGCAAGAGAGAGAGAGTGTGAGTGGG 3540
Db 3103 AAGGTAGTCAACAGGGGCTGAAGAGTGCACAGAGGTACCCATTGTATGGGAATCTG 3162
Qy 3541 CTTACATGGCGATAGCTAGACTGGCGGTTTATGGACAGCAAGCAACCGGAATTGCCA 3600
Db 3163 AT--CTGGGCGCTCGGTGCACATGCTTTACATGTGTTTAGTTCGAGGTTAAAAAGCTTA 3220
Qy 3601 GCTGGGGCGCCTCTGTGAAGTGGGAGCCCTGCAAAAGTAAACTGATGAGCTTTCTTG 3660
Db 3221 GGCCCGCCGAACCAAGGGGAGTGGTTTTCCTTTGAAAAACAACGATGAAGCTTGCCAC 3280
Qy 3661 CCGCCAGGATCTGATGGCGCAGGGATCAAGATCTGATCAAGAGACAGGATGAGGATCG 3720
Db 3281 AACCCGGGAT-----AATTCCTGAGCCAAATATGGGATCG----- 3317
Qy 3721 TTTCGATGATTGAACAGATGATGTCAGCGAGGTTCTCCGGCGCTTGGGTGGAGAGG 3780
Db 3318 -----CCATTGAACAAGATGGATTGACGACAGGTTCTCCGGCGCTTGGGTGGAGAGG 3370
Qy 3781 CTATTCGGCTATGATGGGCAACAACAGACAATCGGCTGCTGTGATGCGCGCTGTTCCGG 3840
Db 3371 CTATTCGGCTATGATGGGCAACAACAGACAATCGGCTGCTGTGATGCGCGCTGTTCCGG 3430
Qy 3841 CTGTCAAGCGAGGGCGCCGGTCTTTTGTCAAGACCGACCTGCGGTGCGCTCAAT 3900
Db 3431 CTGTCAAGCGAGGGCGCCGGTCTTTTGTCAAGACCGACCTGTCGGTGCCTCAAT 3490
Qy 3901 GAACTGACGACGAGGCGCGCTATCGTGGCTGCGACGCGGCGTTCCTTGGCGCA 3960
Db 3491 GAACTGACGACGAGGCGCGCTATCGTGGCTGCGACGCGGCGTTCCTTGGCGCA 3550
Qy 3961 GCTGTGTCGACGTTGTCATGAAGCGGAAAGGACCTGGCTGCTATTTGGGCAAGTGCCTG 4020
Db 3551 GCTGTGTCGACGTTGTCATGAAGCGGAAAGGACCTGGCTGCTATTTGGGCAAGTGCCTG 3610
Qy 4021 GGGCAGGATCTCCTGTCATCTCCTGCTGCTCCGCGAAGATGATCCATCATGCTGAT 4080
Db 3611 GGGCAGGATCTCCTGTCATCTCCTGCTGCTCCGCGAAGATGATCCATCATGCTGAT 3670
Qy 4081 GCAATGGCGGCTGCATACGCTTGTATCCGCTACCTGCGCATTCGACCAACCAAGCGAAA 4140
Db 3671 GCAATGGCGGCTGCATACGCTTGTATCCGCTACCTGCGCATTCGACCAACCAAGCGAAA 3730
Qy 4141 CATGCGATCGAGGACGACGCTACTCGGATGAAAGCGGCTCTGTGTCAGATGATCTG 4200
Db 3731 CATGCGATCGAGGACGACGCTACTCGGATGAAAGCGGCTCTGTGTCAGATGATCTG 3790
Qy 4201 GACCAAGCATCAGGGGCTCGCCGCGCAACTGTTGCGCAGGCTCAAGGCGCGCATG 4260
Db 3791 GACCAAGCATCAGGGGCTCGCCGCGCAACTGTTGCGCAGGCTCAAGGCGCGCATG 3850
Qy 4261 CCCGACGCGGAGGATCTCGTGTGACCCATGGCGATGCTGCTTGGCGCAATATCATGGTG 4320
Db 3851 CCCGACGCGGAGGATCTCGTGTGACCCATGGCGATGCTGCTTGGCGCAATATCATGGTG 3910
Qy 4321 GAAATGGCCGCTTTCTGGATTCATGCACTGTGGCGGCTGGGTGTGGCGGACCGCTAT 4380
Db 3911 GAAATGGCCGCTTTCTGGATTCATGCACTGTGGCGGCTGGGTGTGGCGGACCGCTAT 3970
Qy 4381 CAGGACATAGGTTGGCTACCGGTGATTTGCTGAAGCTTGGCGGCAATGGCTGAC 4440
Db 3971 CAGGACATAGGTTGGCTACCGGTGATTTGCTGAAGCTTGGCGGCAATGGCTGAC 4030
Qy 4441 CGCTTCCTCGTGTGTTAGCGTATCGCGCTCCCGATTCGACGCGCATCGCTTCTATCGC 4500

Db 4031 CGCTTCCTCGTGTGTTACGGTATCGCGCTCCGATTCGACGCGCATCGCTTCTATCGC 4090
Qy 4501 CTTCTTGACGAGTTC-----TTCTG 4520
Db 4091 CTTCTTGACGAGTTCGGTTCGAGCGGATCTGATCAAGAGACAGGATGAGATCGTTTCG 4150
Qy 4521 AGCGGAGCTCGGGTTCGATAAAATAAAAGATTTTATTAGTCTCCAGAAAAAGGGGG 4580
Db 4151 CGCGGAGCTCGGGTTCGATAAAATAAAAGATTTTATTAGTCTCCAGAAAAAGGGGG 4210
Qy 4581 AATGAAAGACCCACCTGTAGTTTGGCAAGCTAGCTTAAGTAACGCAATTTTCAAGGC 4640
Db 4211 AATGAAAGACCCACCTGTAGTTTGGCAAGCTAGCTTAAGTAACGCAATTTTCAAGGC 4270
Qy 4641 ATGGAATAATACATNACTGAGATAGAGATTCAGATCAAGTTCAGGAACAGATGGAAC 4700
Db 4271 ATGGAATAATACATNACTGAGATAGAGATTCAGATCAAGTTCAGGAACAGATGGAAC 4330
Qy 4701 AGCTGAATATGGGCCAAACAGGATATCTGTGTAAGCAGTTCTGTCGCCCGCTCAGGGCC 4760
Db 4331 AGCTGAATATGGGCCAAACAGGATATCTGTGTAAGCAGTTCTGTCGCCCGCTCAGGGCC 4390
Qy 4761 AAGAACAAGTGAACAGCTGAATATATGGGCCAAACAGGATATCTGTGTAAGCAGTTCTG 4820
Db 4391 AAGAACAAGTGAACAGCTGAATATATGGGCCAAACAGGATATCTGTGTAAGCAGTTCTG 4450
Qy 4821 CCCGGCTCAGGGCCAAGAACAGATGGTCCCGAGATGGGTCCAGCCCTCAGCAGTTTCT 4880
Db 4451 CCCGGCTCAGGGCCAAGAACAGATGGTCCCGAGATGGGTCCAGCCCTCAGCAGTTTCT 4510
Qy 4881 AGAACAACATCAGATGTTTCCAGGGTCCCAAGGACCTGAAATGACCCCTGTGCTTATT 4940
Db 4511 AGAACAACATCAGATGTTTCCAGGGTCCCAAGGACCTGAAATGACCCCTGTGCTTATT 4570
Qy 4941 TGAACAAACCAATCAGTTCGCTTCTGCTTCTGTTCCGCGCTTCTGCTCCCCAGCTCA 5000
Db 4571 TGAACAAACCAATCAGTTCGCTTCTGCTTCTGTTCCGCGCTTCTGCTCCCCAGCTCA 4630
Qy 5001 ATAAAGAGGCCCAACACCCCTCCTACCTCGGGCGCCAGTCTCCGATTCAGTTCGCGCCG 5060
Db 4631 ATAAAGAGGCCCAACACCCCTCCTACCTCGGGCGCCAGTCTCCGATTCAGTTCGCGCCG 4690
Qy 5061 GGTACCCGCTGTATCCAATAAACCTCTTGCAGTTGCAATCCGACTTGTGCTCTGCTGTTTC 5120
Db 4691 GGTACCCGCTGTATCCAATAAACCTCTTGCAGTTGCAATCCGACTTGTGCTCTGCTGTTTC 4750
Qy 5121 CTTGGAGGCTCTCTCTGATGATGACTACCCGTGAGCGGGGCTCTTTCATTTGG 5177
Db 4751 CTTGGAGGCTCTCTCTGATGATGACTACCCGTGAGCGGGGCTCTTTCATTTGG 4807

RESULT 9

AR214682 LOCUS 5594 bp DNA linear PAT 25-SEP-2002
DEFINITION Sequence 29 from patent US 6410220.
ACCESSION AR214682
VERSION AR214682.1 GI:23312592
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5594)
AUTHORS Hodgson,C.P., Zink,M.A. and Xu,G.
TITLE Self-assembling genes, vectors and uses thereof
JOURNAL Patent: US 6410220-A 29 25-JUN-2002;
Nature Technology Corp; Omaha, NE
FEATURES
source Location/Qualifiers
1..5594
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 31.5%; Score 1629.4; DB 6; Length 5594;

Db	61	GGAAAAATACATAA	CTGAGATAGAAAAGTT	CAGATCAAGTCTCAGGAACAAGAAACAGC	120
Qy	121	TGAATACCAACACGATATCTGTGTAAAGCGTTT	CCTGCCCGCTCAGGGCCAAAGAACA	180	
Db	121	TGAATACCAACACGATATCTGTGTAAAGCGTTT	CCTGCCCGCTCAGGGCCAAAGAACA	180	
Qy	181	GATGAGACAGCTGAGTGATGGGCCAAACACGATATCTGTGTAAAGCAGTTT	CTCTGCCCCCGG	240	
Db	181	GATGAGACAGCTGAGTGATGGGCCAAACACGATATCTGTGTAAAGCAGTTT	CTCTGCCCCCGG	240	
Qy	241	CTCGGGCCAAAGAACAGATGTTCCCAAGATGCGGTTCAGCCCTCAGCAGTTT	CTTAGTGAA	300	
Db	241	CTCGGGCCAAAGAACAGATGTTCCCAAGATGCGGTTCAGCCCTCAGCAGTTT	CTTAGTGAA	300	
Qy	301	TCATCAGATGTTTCCACGSGTGCCCAAGGACCTGAAATGACCTGTACCTTATTTGAAC	360		
Db	301	TCATCAGATGTTTCCACGSGTGCCCAAGGACCTGAAATGACCTGTACCTTATTTGAAC	360		
Qy	361	TAAACCAATCAGTTCGCTTCTCGCTTCTGTGTCGCGCTTC	CGGTCTCCGAGCTCAATAAA	420	
Db	361	TAAACCAATCAGTTCGCTTCTCGCTTCTGTGTCGCGCTTC	CGGTCTCCGAGCTCAATAAA	420	
Qy	421	AGAGCCACAACCCCTCACTCGGCGCGCAGTCTTCCGATAGACTGCGTCCGCGGATC	480		
Db	421	AGAGCCACAACCCCTCACTCGGCGCGCAGTCTTCCGATAGACTGCGTCCGCGGATC	480		
Qy	481	CCGTATTTCCCAATAAAGCCTCTTCTGTTTGCATCCGAATCGTGGTCTCGCTGTTCCCTTG	540		
Db	481	CCGTATTTCCCAATAAAGCCTCTTCTGTTTGCATCCGAATCGTGGTCTCGCTGTTCCCTTG	540		
Qy	541	GGAGGGTCTCCTCTGAGTGATTGACTACCAACGACGGGGTCTTTCA	TTTGGGGGCTCGT	600	
Db	541	GGAGGGTCTCCTCTGAGTGATTGACTACCAACGACGGGGTCTTTCA	TTTGGGGGCTCGT	600	
Qy	601	CCGGGATTTGGAGACCCCTGCGCCAGGAGCACACCGACCCAC	CCGCGGAGTAAAGCTGGCC	660	
Db	601	CCGGGATTTGGAGACCCCTGCGCCAGGAGCACACCGACCCAC	CCGCGGAGTAAAGCTGGCC	660	
Qy	661	AGCAACTTATCTGTCTGTGTCGATTTCTAGTGTCTATGTTGATGTTATGCGCTCGC	720		
Db	661	AGCAACTTATCTGTCTGTGTCGATTTCTAGTGTCTATGTTGATGTTATGCGCTCGC	720		
Qy	721	TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGCGCGACCCGCTGGTGGAACTGACAGATT	780		
Db	721	TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGCGCGACCCGCTGGTGGAACTGACAGATT	780		
Qy	781	CTGAACACCCGGCCGCAACCTCTGGGAGACGTTCCAGGGACCTTTTGGGGCCGTTTTGTGG	840		
Db	781	CTGAACACCCGGCCGCAACCTCTGGGAGACGTTCCAGGGACCTTTTGGGGCCGTTTTGTGG	840		
Qy	841	CCCGACTGTAGGAAGGAGTGCATGTGGAAATCCGACCCCGTCAAGATATGTGTTCTGGT	900		
Db	841	CCCGACTGTAGGAAGGAGTGCATGTGGAAATCCGACCCCGTCAAGATATGTGTTCTGGT	900		
Qy	901	AGGAGACGAGAACCTAAACAGTTTCCCGCTCTGATTTTGGCTTTCCGTTTGGAA	960		
Db	901	AGGAGACGAGAACCTAAACAGTTTCCCGCTCTGATTTTGGCTTTCCGTTTGGAA	960		
Qy	961	CCGAAGCCGCGCTTGTCTGTCTGCGACCAAGCTTGGGGTGCAGGTCGACTCTAGAGGA	1020		
Db	961	CCGAAGCCGCGCTTGTCTGTCTGCGACCAAGCTTGGGGTGCAGGTCGAGGACTGGGA	1020		
Qy	1021	TCAATTCCGCCACGAGTAAATCGGTGTCGCGTCTTTAGGACATATGAAGTATGGCAAGT	1080		
Db	1021	CC-----CTGCACCGAAATGGAGAACACATCATCAGGATTCCT	1059		
Qy	1081	GGGATGACTTTCGTGATCAGCAAGAGAGACATGACAGCTGTACAGAGCTGTGAGAGTTCC	1140		
Db	1060	AGGACCCCTGCTGTGTTACAGCGGGGTTTTCTTGTGACAAAGAAATCTCACAATACC	1119		
Qy	1141	ATGCTCGCTCAGTGACAGCTTTGTCTCTCCCATCTCTAAATAATGGGCCAACTCTTCAAG	1200		

2281 ACBAAGTCTCTCTGGACTCCAGTGTGAAGGAGATAGAGGCCCTCTCTGACAAAATGG 2340
2006 AACAAATATCTGAACCTTTTACCCCGTTGCCGGCA----- 2039
2341 TATACAGGCTTTCAGGCTCTAATAGTACTCCAGGCTTTAAAGGTGATCGGGGGATCT 2400
2040 ----- 2039
2401 CTGGTTTACCTGGAGTTGAGGATTTCCAGGACCAATGGGGAGACCGGGAAGCCAGGAC 2460
2040 ----- 2039
2461 TTAATGACAAAAGGCCAGAGAGGAGAAAAGGGAGTGGAGCATGCAAGACAAATCTTA 2520
2040 -----ACGGTCAG 2047
2521 ATACAGTCCGACTGGTGGGTGGCAGCGCCCTCACGAAGCAGAGTGGAGATTTTTCACG 2580
2048 GTCTCTCCAAAGTGTCTGACGCAACCCCACTGGATGGGCTTGGCTATCGGCCATA 2107
2581 AAGGCCAGTGGGTGACGCTGTGTGACGACCGCTGGGAACTGCGTGGAGGACTGTCGTCT 2640
2108 GCCGATGCGCGACCTTTGTGGCTCTCTGCCG----- 2141
2641 GCAGGAGCTTGGGATACAAAGGTGTTCAAAGTGTGCAATAGCAGCTTATTTGGAAAAG 2700
2142 -----ATCCATATCGGAACTCTCTAGCAGCTTGT----- 2171
2701 GTACGGGTCCAAATATGCTGAATGAAGTATTTGTTTCGGGAAGAGTCACTCATTTGAAG 2760
2172 ----- 2171
2761 AGTGCAGAAATTAGACAGTGGGTGTGAGAGCTGTTTCGCACGACGAAGATGCTGGGGTC 2820
2172 -----TTTGCTCGCAGCGGCTCGAGCGAA 2197
2821 ACTTTGCACTACATAATGATCATATTTTTCATTCACATTTTAAACTGTTATAAGTG 2880
2198 ACTTATCGGC----- 2207
2881 ATTTTTTCTTTCCTTCACTAAATCAGCTTAATTAATTTAAGAAACTAAGAAATTTT 2940
2208 ----- 2207
2941 ATCCACAGAAAAGGAATTTTAAATAATCACTGGATAAACAATAATAAATAGCTTCATATTT 3000
2208 -----ACCGAACACTCT 2219
3001 GCTTCAAATACCAAGAACATTTCAACTTCTCTAGGTTTTTAAAGTGGCTCGTGCCGAATTG 3060
2220 GTTGTCTCTCTCGGAAATACACCTCTTTTCCATGGCTGTAGGCTGTGCTGCCAATGG 2279
3061 ATCCCTCAGGATATAGTAGTTGCTTTTTCATAGGAGGGGGAATAGTAGTCTTATGC 3120
2280 ATCCCTCAGGATATAGTAGTTGCTTTTTCATAGGAGGGGGAATAGTAGTCTTATGC 2339
3121 AATACTCTCTAGTCTTGAACATGTAACGATGAGTTAGCAATGCTCCCTTACAGGAGA 3180
2340 AATACACTTGTAGTCTTGAACATGTAACGATGAGTTAGCAATGCTCCCTTACAGGAGA 2399
3181 GAAAAGCACCGTGCATGCGATGTTGTTGGAAGTAAAGTGTAGCATCGTCCCTTATTAGG 3240
2400 GAAAAGCACCGTGCATGCGATGTTGTTGGAAGTAAAGTGTAGCATCGTCCCTTATTAGG 2459
3241 AAGCAACACAGCGGTCTGATGATGATGAGTGAAGCAACCACTGAATTCGCAATTCAGAGAT 3300
2460 AAGCAACACAGAGGTCTGATGATGATGAGTGAAGCAACCACTGAATTCGCAATTCAGAGAT 2519
3301 -ATTGATTTTAAAGTCTCTGATACAGCAACGCA--TTTGACCATTTCCACCATTT 3357
2520 AATTGATTTTAAAGTCTCTGATACAGCAACGCAATTTTGAACCATTTCCACCATTT 2579

3358 GGTGTGACCT--CCAAGCTTCACTGCTCCGCAAGCACTCAGGGCGCAAGGGCTGCTAAA 3415
2580 GGTGTGACCTTCCAAAGCTTCACTGCTCCGCAAGCACTCAGGGCGCAAGGGCTGCTAAA 2639
3416 GGAAGCGGAACACGTAGAAAGCAGTCCGAGAAAAGGCTGTGACCCCGGATGAATGTCA 3475
2640 GGAAGCGGAACACGTAGAAAGCAGTCCGAGAAAAGGCTGTGACCCCGGATGAATGTCA 2699
3476 GCTACTGGGCTATCTGGCAAGGGGAAAACGCAAGCGCAAAAGAGACAGGTAGCTTGA 3535
2700 GCTACTGGGCTATCTGGCAAGGGGAAAACGCAAGCGCAAAAGAGACAGGTAGCTTGA 2759
3536 GTGGCTTACATCGCGATAGTAGACTGGCGGTTTTTATGACAGCAAGCAAGCGAAT 3595
2760 GTGGCTTACATCGCGATAGTAGACTGGCGGTTTTTATGACAGCAAGCAAGCGAAT 2819
3596 TGCAGCTGGGGCGCCCTCTGTTAAAGTTGGGAAGCCCTGCAAGTAAACTGATGCTT 3655
2820 TGCAGCTGGGGCGCCCTCTGTTAAAGTTGGGAAGCCCTGCAAGTAAACTGATGCTT 2879
3656 TCTTGGCGCAAGGATCTGATGGCGAGGGGATCAAGATCTGATCAAGAGACAGGATGAG 3715
2880 TCTTGGCGCAAGGATCTGATGGCGAGGGGATCAAGATCTGATCAAGAGACAGGATGAG 2939
3716 GATCGTTTGCATGATGAAAGATGGATGGACAGGATTTCTCCGGCCGCTTGGGTGG 3775
2940 GATCGTTTGCATGATGAAAGATGGATGGACAGGATTTCTCCGGCCGCTTGGGTGG 2999
3776 AGAGGCTATTCGGCTATGACTGGGCAACAGCAAACTCGGCTGCTGTGATGCGGCCGTGT 3835
3000 AGAGGCTATTCGGCTATGACTGGGCAACAGCAAACTCGGCTGCTGTGATGCGGCCGTGT 3059
3836 TCCGGCTGTGAGCGCGAGGGCGCCGGTTCTTTTGTCAAGACCGACTCTGCTCCGGTGCCC 3895
3060 TCCGGCTGTGAGCGAGGGGCGCCGGTTCTTTTGTCAAGACCGACTCTGCTCCGGTGCCC 3119
3896 TGAATGAATTCGACGAGCGAGCGGCTATGTTGGCTGGCCACGACGGGGCTTCCCTT 3955
3120 TGAATGAATTCGACGAGCGAGCGGCTATGTTGGCTGGCCACGACGGGGCTTCCCTT 3179
3956 GCGACGTGCTCGACGTTGTCATGAGCGGGAAGGACTGGCTGCTTATTTGGGCGAAG 4015
3180 GCGACGTGCTCGACGTTGTCATGAGCGGGAAGGACTGGCTGCTTATTTGGGCGAAG 3239
4016 TGCCGGGCGAGGATCTCTGTCTATCTCACCTTCTCTGCGGAGAAAGTATTCATCATGG 4075
3240 TGCCGGGCGAGGATCTCTGTCTATCTCACCTTCTCTGCGGAGAAAGTATTCATCATGG 3299
4076 CTGATGCAATGCGCGCTGCATACGCTTGTATCCGGCTACCTGCCATTCGACCAACG 4135
3300 CTGATGCAATGCGCGCTGCATACGCTTGTATCCGGCTACCTGCCATTCGACCAACG 3359
4136 CGAAATCTGCAATGCGCGCTGCATACGCTTGTATCCGGCTACCTGCCATTCGACCAACG 4195
3360 CGAAATCTGCAATGCGCGCTGCATACGCTTGTATCCGGCTACCTGCCATTCGACCAACG 3419
4196 ATCTGACCAAGAGCATCAGGGCTCGCGACCGCAACTGTTTCGCGAGGCTCAAGCGC 4255
3420 ATCTGACCAAGAGCATCAGGGCTCGCGACCGCAACTGTTTCGCGAGGCTCAAGCGC 3479
4256 GCATGCCCGACGCGAGGATCTCGTGTGACCCATGCGATGCTGCTGCCGAATATCA 4315
3480 GCATGCCCGACGCGAGGATCTCGTGTGACCCATGCGATGCTGCTGCCGAATATCA 3539
4316 TGGTGAATAATGCGCGCTTCTGATTCATTCGACTGTGCGCGCTGGGTGTGCGGAC 4375
3540 TGGTGAATAATGCGCGCTTCTGATTCATTCGACTGTGCGCGCTGGGTGTGCGGAC 3599
4376 GCTATCAGGACATAGGCTTGGCTACCGCTGATATTCGTAAGAGCTTGGCGGGAATGG 4435
3600 GCTATCAGGACATAGGCTTGGCTACCGCTGATATTCGTAAGAGCTTGGCGGGAATGG 3659
4436 CTGACCGCTTCTCTGCTGCTTTTACGATGCGCGCTCCCGATTCGACGCGCATGCTCT 4495

|||||
Db 3660 CTGACCGCTTCTCGTGTCTTTAGGGTATCGCGCTCCGGATTCGACGGCATCGCTTCT 3719
Qy 4496 ATCGCCTTCTTGACGAGTCTTCTGACGGGACTCTGGGTTGATA 4542
Db 3720 ATCGCCTTCTTGACGAGTCTTCTGACGGGACTCTGGGTTGAAA 3766
RESULT 11
AX359935
LOCUS AX359935 5130 bp DNA linear PAT 13-FEB-2002
DEFINITION Sequence 9 from Patent WO020783.
ACCESSION AX359935
VERSION AX359935.1 GI:18675574
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; 1
REFERENCE 1
AUTHORS Bleck, G.T.
TITLE Expression vectors
JOURNAL Patent: WO 020783-A 9 10-JAN-2002;
Gala Design, Inc. (US)
FEATURES
source 1..5130
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic"
ORIGIN
Query Match 30.78; Score 1588.6; DB 6; Length 5130;
Best Local Similarity 66.18; Pred. No. 0;
Matches 3007; Conservative 0; Mismatches 754; Indels 786; Gaps 18;
Qy 1 TTTGAAAGACCCCAACCGTAGGTGGCAAGCTAGCTTAAGTAAACGCCACTTTTGAAGGCAT 60
Db 1 TTTGAAAGACCCCAACCGTAGGTGGCAAGCTAGCTTAAGTAAACGCCACTTTTGAAGGCAT 60
Qy 61 GGAATAATACATACTAGATAAGAAAGTTTCAGATCAAGGTCAAGAAACAAAGAAACAGC 120
Db 61 GGAATAATACATACTAGATAAGAAAGTTTCAGATCAAGGTCAAGAAACAAAGAAACAGC 120
Qy 121 TGATACCAACAGGATATCTGGTAAGCGGTTCTCGCCCGCTCAGGGCCAAAGAA 180
Db 121 TGAATACCAACAGGATATCTGGTAAGCGGTTCTCGCCCGCTCAGGGCCAAAGAA 180
Qy 181 GATGAGACAGCTGAGTGATGGGCCAAACAGGATATCTGGTAAGCGGTTCTCGCCCGG 240
Db 181 GATGAGACAGCTGAGTGATGGGCCAAACAGGATATCTGGTAAGCGGTTCTCGCCCGG 240
Qy 241 CTCGGGGCCAAAGAACAGATGTCCCAAGATGCGGTCCAGCCCTCAGCGATTCTAGTGAA 300
Db 241 CTCGGGGCCAAAGAACAGATGTCCCAAGATGCGGTCCAGCCCTCAGCGATTCTAGTGAA 300
Qy 301 TCATCAGATCTTCCAGGTGCCCCAAGGACCTGAAATGACCTGTACCTTATTTGAAC 360
Db 301 TCATCAGATGTCTCCAGGTGCCCCAAGGACCTGAAATGACCTGTACCTTATTTGAAC 360
Qy 361 TAAACCAATCAGTTCGCTTCTCGCTTCTGTTGGCGGCTTCGGCTCTCCGAGCTCAATAAA 420
Db 361 TAAACCAATCAGTTCGCTTCTCGCTTCTGTTGGCGGCTTCGGCTCTCCGAGCTCAATAAA 420
Qy 421 AGAGCCCAACCCCTCACTCGGGCGCCAGTCTTCGATAGACTGCGTCCCGGGTAC 480
Db 421 AGAGCCCAACCCCTCACTCGGGCGCCAGTCTTCGATAGACTGCGTCCCGGGTAC 480
Qy 481 CCGTATTTCCCAATAAAGCCTTGTCTTTGATCCGAAATCGTGGTCTCGCTGTTCTTGG 540
Db 481 CCGTATTTCCCAATAAAGCCTTGTCTTTGATCCGAAATCGTGGTCTCGCTGTTCTTGG 540
Qy 541 GGAGGGTCTCTCTGAGTGATTTGACTACCAACAGACGGGGTCTTTTCATTTGGGGGCTCGT 600
|||||

Db 541 GGAGGGTCTCTCTGAGTGATTGACTACCCACGACGGGGTCTTTTCATTTGGGGGCTCGT 600
Qy 601 CCGGGATTTGGAGACCCCTGCCAGAGACCCACGACCCACCGGGAGGTAGAGCTGGCC 660
Db 601 CCGGGATTTGGAGACCCCTGCCAGAGACCCACGACCCACCGGGAGGTAGAGCTGGCC 660
Qy 661 AGCAACTTATCTGTCTGTCTGTCCGATTTGTCTAGTGTCTATGTTTATGCGCCTGCG 720
Db 661 AGCAACTTATCTGTCTGTCTGTCCGATTTGTCTAGTGTCTATGTTTATGCGCCTGCG 720
Qy 721 TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGACCCGTTGGTGAACCTGACAGTT 780
Db 721 TCTGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGACCCGTTGGTGAACCTGACAGTT 780
Qy 781 CTGAACACCCGGCCGCAACCTCGGAGACGTCCTCAGGACCTTTGGGGGCGCTTTTGTGG 840
Db 781 CTGAACACCCGGCCGCAACCTCGGAGACGTCCTCAGGACCTTTGGGGGCGCTTTTGTGG 840
Qy 841 CCGGACCTGAGGAAGGAGTTCGATGTGGAATCCGACCCCGTCAGGATATGTGTTCTGTT 900
Db 841 CCGGACCTGAGGAAGGAGTTCGATGTGGAATCCGACCCCGTCAGGATATGTGTTCTGTT 900
Qy 901 AGGAGACGAGAACCTAAACAGTTCCGCTCCGCTCTGAATTTTTCGTTTCGTTGGAA 960
Db 901 AGGAGACGAGAACCTAAACAGTTCCGCTCCGCTCTGAATTTTTCGTTTCGTTGGAA 960
Qy 961 CCGAAGCCGCGCTTGTCTGCTGAGCCCAAGCTTTGGGCTGAGGACTGAGGAA 1020
Db 961 CCGAAGCCGCGCTTGTCTGCTGAGCCCAAGCTTTGGGCTGAGGACTGAGGAA 1020
Qy 1021 TCAATTCGGCACGAGTAAATCGGTGCTCTTAGGACATATGAAGTATGACACAGT 1080
Db 1021 CC-----CTGCACCGAATCGGAAACACACATCAGGATTCCT 1059
Qy 1081 GGGATGACTTCTGTATCAGCAAGAGGACACTGACAGCTCTAAGAGTCTGTGAAGTTG 1140
Db 1060 AGGACCCCTGCTCGTGTACAGGCGGGTCTTTCTGTTGACAAGAACTCTCAATACC 1119
Qy 1141 ATGCTCGCTCAGTGACAGCTTTGCTTCTCCCTCCCATCTTAAATGCCCCAATCTTCAAG 1200
Db 1120 ACAGAGTCTAGACTCGTGTGGACTTCTCTCAATTTTCTAGGGGAGACCCACGTTCC 1179
Qy 1201 AGAGGATGAAGTCTTATAAAATGCACTGATCACTCTTATCTCATTTGTTGTAGTTTC 1260
Db 1180 TGCCCAAAATTCGAGTCCCAACCTCCATCACTCAACCACTCTGTGCTCCAATTTG 1239
Qy 1261 TCGTCCCATCATTTGGCATAGTGGCAGCTCAGCTCTCTGAATCGGAAACGAGAAATTTGCA 1320
Db 1240 TCTGGCTATC-----GCTGGATGTCTCTGCGCGCTTTTATC 1276
Qy 1321 CGTTGGCTCAGTTAATGAGATATATCTCCAAGTCCGNAAGGCAAGAAATGGCAGTG 1380
Db 1277 ATATTCCTCTTCACTCTGCTGCTATGCTCATCTTCTGTTGTTCTTCTGGACTACAA 1336
Qy 1381 AAGATGAATAGATTTTCGAGAGAGTGTGATGAAACGATGAGCAACATCGAAAGCAGAA 1440
Db 1337 GGTATGTGGCCGTTTGTCTCTACTTCCAGGAA---CATCACTACAGACGCGGACCA 1393
Qy 1441 TCAGATATCTTTCAGATAATGAAGCCAAATCTCTAGATGCTAAGAAATTTCCAAATTTCA 1500
Db 1394 TGCAAGACCTGCAAGATTCCTGTCTCAAGGAACTCTATGTTTCCCTCTT-----GTTGCT 1448
Qy 1501 GATATACACTGATCAAGATTTAATGATGTTCTTTTCCAGCTTAATTTCTTACTTTCCT 1560
Db 1449 GTACAAAACCTTCGGACGGAACTGCATTTGTATTCCTATCCCATCATCTCGGGCTTTG 1508
Qy 1561 CCATCCAGGAACATGAGATATCATAGGGGATATCTCCAAGTCAATAGTAGTCTGAACA 1620
Db 1509 CAAGATTCCTATGGGAGTGGGCTCAGTCCGTTTCTCTCG----- 1549
Qy 1621 CCAAGTACTTGTATTTGCGAGTTCCAGTATTGAAACACTGAAATGCGCAGAGTCCAGAGAA 1680
Db 1550 -CTCAGTTTACTAGTGCCATTTTGTTCAGTGGTTC---GTAGGGCTTTTCCCGCCTGTTG 1605
|||||

Qy	1691	CATTAAACAACAAGAGAGATGCGTAAATTTAGAGGACGCTATATACATGCATCAGCAG	1741
Db	1606	GCTTTTCAGTTATATG-GAATGAATGGTATTCGGGGCCAAAGTCTCTGACAAATCTTTGATC	1664
Qy	1741	AAATTAAGTCTCTAGATGAAAAACAAGTATATTTGGAAACAGGAAATAAAAGGGGAAATGA	1800
Db	1665	CCTTTTACCTCTATTACCAATTTTCTTTGTCTTTGGGTATACATTTAAACCTTATAA	1724
Qy	1801	AACGTGTTGAATAATATCACTAATGATCTGAGGCTGAAGGATTTGGGAACATTTCTCAGACAT	1860
Db	1725	AAC-----CAAAAGT-----CAAAAGT-----CAAAAGT-----CAAAAGT-----	1734
Qy	1861	TGAAAAATATCACTTTACTTCAAGTGCAGAAAGTCTCGCTGACTGGGAAATGGACCA	1920
Db	1735	TGGGGTACTCCCTTAATCTCATCGGGATATGTAATGGATTTGGGG-----	1781
Qy	1921	ACGATCTGGGCTCCAACATGACCATCGGGGCTGTGAAACAGCAGAGGTGAATTTACAGGCA	1980
Db	1782	-----	1781
Qy	1981	CCTACATCAGACGCGTAACGCCACATCAATGAGATCAAAAGATCAACACTGATGGGA	2040
Db	1782	-----TACTTTACGCCAAGAACATATTGTACTAAAAATCAACAGCAATGTTTTTCG-	1829
Qy	2041	CACAAAACACCATCAACAAGAGGACCCAGCCACCTTTTGGCTTCACCGTCAATTTGGAAGT	2100
Db	1830	-----AAACTGCTGTAAATAGACCTATTGATTTGGAAAGTATGTCAGAGACTTTGTGGT	1884
Qy	2101	TTTCAGAGTCCACCACTGTCTTTCACGGGCCAGTCTTTCATAGACAGAAATGGGAAGG	2160
Db	1885	CTTTTGGGCTTTGTGCCCCCTTTTACACAATGTGGCTATCCTGCCTTAATGCCCCTTATAT	1944
Qy	2161	TCCTGAAGACCATGTGGCTGTCGGGTCAAGTGTTAATGATCATTTGTGTGATGACTGGAAAG	2220
Db	1945	GCATGTATACAATCT-----	1959
Qy	2221	CTACGAGGTGGGCATCAACATCTTCACTCGCTCGCCGACACACAGAGGAGTGAAGTGCAGTG	2280
Db	1960	-----AAGCAGGCTTTCACTTTCTCGCCACTTACAGAGCCTTTCTGTGTGA	2005
Qy	2281	ACCAAGGTCTCTCTGGACTCCAGGTGAAAAAGAGATAGAGGCCCTCTCGACAAAAATGG	2340
Db	2006	AACAAATATCGAACCTTTTACCCGTTGCCGGCA-----	2039
Qy	2341	TATACAGGCTTTCCAGGTCTAATAGTACTCCAGGTCTTAAAGGTGATCGGGGGGATCT	2400
Db	2040	-----	2039
Qy	2401	CTGTTTACCTGGAGTTCGAGATTCCACGACCATTTCCAGGAGACCGGGAAGCCAGGAC	2460
Db	2040	-----	2039
Qy	2461	TTAATGGCAAAAAGGCCAGAGGGAGAAAAGGAGTGAAGCATGCAAAAGCAATCTA	2520
Db	2040	-----ACGGTCAG	2047
Qy	2521	ATACAGTCCGACTGGTGGGTGGCAGCGGCCTCACGAAGGAGATGGAGATTTTTTCACG	2580
Db	2048	GTCTCTCCCAAGTGTGTTGCTGAACCAACCCCACTGGATGGGGCTTGGCTATCGGCCATA	2107
Qy	2581	AAGGCCAGTGGGTACGGTGTGTGACACACCGCTGGGAACCTCGCTGGAGGACTGTCGTCT	2640
Db	2108	GCCGCATGCGCGAACCTTTGTGGCTCTCTGCGG-----	2141
Qy	2641	GCAGGAGCTTTGGGATACAAAGGTGTTCAAAGTGTGCATAAGCGAGCTTATTTTGGAAAAG	2700
Db	2142	-----ATCCATACTGGGAACCTCCTAGCAGCTTG-----	2171
Qy	2701	GTACGGGTCCATATAGTGCTGAATGAAGTATTTGTTTCGGGAAAGAGTCAATCAATTGAAG	2760
Db	2172	-----	2171

Qy	2761	AGTGCAGAA	TTAGCAGCTGGGGTGTGAGAGCCCTGTTTCGCACGACGAAAGATGCTGGGGGTC	2822	
Db	2172	-----	TTTGTGCGACGCGCTCTGGAGCGAA	2197	
Qy	2821	ACTTTGACCTACATAATATGATCATATTTTCA	TTTCAATTTTTTAAACTGTTTATAAAGTG	2880	
Db	2198	ACTTATCGGC	-----	2207	
Qy	2881	ATTTTTTTCCTTGGCTTCACTAAATACGCTTAATTA	TTTAAGAAACTAAGAATTTT	2940	
Db	2208	-----	-----	2207	
Qy	2941	ATCCACAGAAAAGGAATATTTAAAAATCACTGATAAACATATATAAAATAGCTTCATATTT	3000		
Db	2208	-----	-----ACCACCACTCT	2219	
Qy	3001	GCTTCAAAATACCAGAACCA	TTTCAA	CTTCTAGGTTTTTAAAGTGCTCGTCCGAATG	3060
Db	2220	GTGTCTCTCTCGGAAATACACTCTT	TTCCATGCTGCTAGGTGTCTGCCA	CTGG	2279
Qy	3061	ATCCCCTCAGGATATAGTAGTTTCGTTTTG	CNATAGGAGGGGAAATAGTCTTTATGC	3120	
Db	2280	ATCCCCTCAGGATATAGTAGTTTCGTTTTG	CNATAGGAGGGGAAATAGTCTTTATGC	2339	
Qy	3121	AATACTCTTGTAGTCTTGCNACATGTGTA	CGATGAGTTAGCAACATGCCTTACAGGAGA	3180	
Db	2340	AATACACTTGTAGTCTTGCNACATGTGTA	CGATGAGTTAGCAACATGCCTTACAGGAGA	2399	
Qy	3181	GA AAAAGCACCGTGCATCCGATTG	TGTGGAAGTAAAGTGCTACGATCGTGCCCTTATTAGG	3240	
Db	2400	GA AAAAGCACCGTGCATCCGATTG	TGTGGAAGTAAAGTGCTAGTCTGTCCTTTATTAGG	2459	
Qy	3241	AAGGCAACAGACGGGCTCTGACATG	SGATTTGGAAGCAACCACTGAAATCCCGCA	TTTCGAGAGAT	3300
Db	2460	AAGGCAACAGACAGGCTCTGACATG	SGATTTGGAAGCAACCACTGAAATCCCGCA	TTTCGAGAGAT	2519
Qy	3301	ATTGTATTTAAGTGCTAGCTCGATACAG	CAAAACGCCA - TTTGACCAATTCACCA	CATTT	3357
Db	2520	AATTTGATTTAAGTGCTAGCTCGATACAG	CAAAACGCCA TTTTGTGACCAATTCACCA	CATTT	2579
Qy	3358	GGTGTGCACCT - CCAAGCTTCA	CGCTGCCGCAAGCACTCAGGCGCAAGGGCTGCTAAA	3415	
Db	2580	GGTGTGCACCTTCCAAGCTTCA	CGCTGCCGCAAGCACTCAGGCGCAAGGGCTGCTAAA	2639	
Qy	3416	GGAAAGCGGAACAGTATAGAAAGCCAGTCCG	CAGAAAACGGTGTCTGCACCCCGGATGAATGTCA	3475	
Db	2640	GGAAAGCGGAACAGTATAGAAAGCCAGTCCG	CAGAAAACGGTGTCTGCACCCCGGATGAATGTCA	2699	
Qy	3476	GCTACTGGGCTATCTCGACAAGGGGAAAA	CGCAAGCGCAAAAGAGCAAGTAGCTTGCA	3535	
Db	2700	GCTACTGGGCTATCTCGACAAGGGGAAAA	CGCAAGCGCAAAAGAGCAAGTAGCTTGCA	2759	
Qy	3536	GTGGGCTTACATGGCCATAGCTAGACTGG	CGGGTTTTATGGACACGACGCAACCGGAT	3595	
Db	2760	GTGGGCTTACATGGCCATAGCTAGACTGG	CGGGTTTTATGGACACGACGCAACCGGAT	2819	
Qy	3596	TGCGAGCTGGGCGGCCCTCTGGTAA	GGTGGGAAGCCCTGCAAAAGTAAACTGGA	TGGCTT	3655
Db	2820	TGCGAGCTGGGCGGCCCTCTGGTAA	GGTGGGAAGCCCTGCAAAAGTAAACTGGA	TGGCTT	2879
Qy	3656	TCTTTGCGGCAAGGATCTGATGCGCAG	GGGATCAAGATCTCATCAAGAGACAGGATGAG	3715	
Db	2880	TCTTTGCGGCAAGGATCTGATGCGCAG	GGGATCAAGATCTCATCAAGAGACAGGATGAG	2939	
Qy	3716	GATCGTTTCGCATGATTTGAACAAGATG	GAATTCGACGAGGTTCTCGGCGGCTTGGTG	3775	
Db	2940	GATCGTTTCGCATGATTTGAACAAGATG	GAATTCGACGAGGTTCTCGGCGGCTTGGTG	2999	
Qy	3776	AGAGGCTATTTCGGCTATGACTGGGCA	CAACAGACAAATCGCTGCTCTGATGCGCGCTG	3835	
Db	3000	AGAGGCTATTTCGGCTATGACTGGGCA	CAACAGACAAATCGCTGCTCTGATGCGCGCTG	3059	
Qy	3836	TCCGGCTGTGACGCGCAGGGGCGCC	CGGTTCTTTTTTGTCAAGACCGACCTGTCCG	TGCC	3895

Db 3060 TCCGGCTGTCAGCGCAGGGCGCCGGTCTCTTTTGTCAAGACCGACCTGTCGGGTGCC 3119
Qy 3896 TGAATGAATCTCAGACGAGGCGCGCTATCGTGGCTGCGCCACGAGCGGCGTTCCTT 3955
Db 3120 TGAATGAATCTCAGACGAGGCGCGCTATCGTGGCTGCGCCACGAGCGGCGTTCCTT 3179
Qy 3956 CGCAGCTGTCTCGACGCTTCTCACTGAAGCGGGAAGGGAATGGCTGCTATTTGGCGGAAG 4015
Db 3180 GCGCAGCTGTCTCGACGCTTCTCACTGAAGCGGGAAGGGAATGGCTGCTATTTGGCGGAAG 3239
Qy 4016 TGGCGGGCAGGATCTCTGTCTATCTCACTTGTCTGCTGCGGGAAGGGAATGGCTGCTATTTGGCGGAAG 4075
Db 3240 TGGCGGGCAGGATCTCTGTCTATCTCACTTGTCTGCTGCGGGAAGGGAATGGCTGCTATTTGGCGGAAG 3299
Qy 4076 CTGATGCAATCGCGGGCTGATAGCTTGTGATCGGCTACCTGCGGCTTGCACCAACCAAG 4135
Db 3300 CTGATGCAATCGCGGGCTGATAGCTTGTGATCGGCTACCTGCGGCTTGCACCAACCAAG 3359
Qy 4136 CGAAACATCGCATCGAGCGAGCAGTACTCGGATGGAAGCGGCTTGTGATCAGGATG 4195
Db 3360 CGAAACATCGCATCGAGCGAGCAGTACTCGGATGGAAGCGGCTTGTGATCAGGATG 3419
Qy 4196 ATCTGGAACGAGAGCATCAGGGGCTCGCGCAGCGGAACTGTTTCCGAGGCTCAAGCGC 4255
Db 3420 ATCTGGAACGAGAGCATCAGGGGCTCGCGCAGCGGAACTGTTTCCGAGGCTCAAGCGC 3479
Qy 4256 GCATGCCCGAGCGGAGATCTCTGATGAGCCATGCGGATGCGGCTTGCAGGATATCA 4315
Db 3480 GCATGCCCGAGCGGAGATCTCTGATGAGCCATGCGGATGCGGCTTGCAGGATATCA 3539
Qy 4316 TGGTGGAAATGGCGGCTTTTCTGATTCATCGACTGTGGCGGCTGGGTGTCGGGACC 4375
Db 3540 TGGTGGAAATGGCGGCTTTTCTGATTCATCGACTGTGGCGGCTGGGTGTCGGGACC 3599
Qy 4376 GCTATCAGGACATAGCGTGTGCTACCGGTGATATGCTGAAGAGCTTGGCGGGAATGGG 4435
Db 3600 GCTATCAGGACATAGCGTGTGCTACCGGTGATATGCTGAAGAGCTTGGCGGGAATGGG 3659
Qy 4436 CTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTCGAGGCAATGCCCTTCT 4495
Db 3660 CTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTCGAGGCAATGCCCTTCT 3719
Qy 4496 ATCGCTTCTTGACGAGTCTTCTGAGCGGACTCTGGGGTTCGATA 4542
Db 3720 ATCGCTTCTTGACGAGTCTTCTGAGCGGACTCTGGGGTTCGATA 3766

RESULT 12
AX382149 5130 bp DNA linear PAT 18-MAR-2002
LOCUS
DEFINITION Sequence 9 from Patent WO0202738.
ACCESSION AX382149
VERSION AX382149.1 GI:19576958
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Bremel,R.D., Miller,L.U., Bleck,G.T. and York,D.
TITLE Host cells containing multiple integrating vectors
JOURNAL Patent: WO 0202738-A 9 10-JAN-2002;
Gala Design, Inc. (US)
FEATURES
source
1. 5130
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

ORIGIN

Query Match 30.7%; Score 1588.6; DB 6; Length 5130;
Best Local Similarity 66.1%; Pred. No. 0;
Matches 3007; Conservative 0; Mismatches 754; Indels 786; Gaps 18;

Qy 1 TTTGAAAGACCCCAACCGTAGGTGGCAAGCTAGCTTAAGTAAACGCCACTTTTGCAGGCAT 60
Db 1 TTTGAAAGACCCCAACCGTAGGTGGCAAGCTAGCTTAAGTAAACGCCACTTTTGCAGGCAT 60
Qy 61 GGAATAATACATTAATGAGAAATAGAAAGTTTCAGATCAAGGTCAAGGAACAAGAAACAGC 120
Db 61 GGAATAATACATTAATGAGAAATAGAAAGTTTCAGATCAAGGTCAAGGAACAAGAAACAGC 120
Qy 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTCCTGCCCCGGCTCAGGGCCCAAGAAC 180
Db 121 TGAATACCAACAGGATATCTGTGTAAGCGGTTCCTGCCCCGGCTCAGGGCCCAAGAAC 180
Qy 181 GATGAGACAGCTGATGATGGGCAACAGGATATCTGTGTAAGCGGTTCCTGCCCCGG 240
Db 181 GATGAGACAGCTGATGATGGGCAACAGGATATCTGTGTAAGCGGTTCCTGCCCCGG 240
Qy 241 CTCGGGCGCAAGAACAGATGTTCCCAAGTCCCGGTCAGGCTCAGAGTTCCTAGTGA 300
Db 241 CTCGGGCGCAAGAACAGATGTTCCCAAGTCCCGGTCAGGCTCAGAGTTCCTAGTGA 300
Qy 301 TCATCAGATGTTTCCAGGGTCCCAAGGACCTGAAATGACCTGTACCTTATTTGAAC 360
Db 301 TCATCAGATGTTTCCAGGGTCCCAAGGACCTGAAATGACCTGTACCTTATTTGAAC 360
Qy 361 TAACCAATCAGTTCGGCTTCTGCTTCTGTTGCGCGCTTCCGGCTCTCCGAGCTCAATAA 420
Db 361 TAACCAATCAGTTCGGCTTCTGCTTCTGTTGCGCGCTTCCGGCTCTCCGAGCTCAATAA 420
Qy 421 AGAGCCCAACCCCTCACTCGCGCGCCAGTCTTCCGATAGACTGCGTCCCGCGGTAC 480
Db 421 AGAGCCCAACCCCTCACTCGCGCGCCAGTCTTCCGATAGACTGCGTCCCGCGGTAC 480
Qy 481 CCGTATTCGCAATAAAGCCTTCTGCTGTTTGCATCCGAATCGTGGTCTCCCTGTTCTTG 540
Db 481 CCGTATTCGCAATAAAGCCTTCTGCTGTTTGCATCCGAATCGTGGTCTCCCTGTTCTTG 540
Qy 541 GGAAGGTCTCTCTGATGATGACTACCCACGAGGGGTCTTTTCAATTTGGGGGTCTGT 600
Db 541 GGAAGGTCTCTCTGATGATGACTACCCACGAGGGGTCTTTTCAATTTGGGGGTCTGT 600
Qy 601 CCGGGAATTTGGAGACCCCTGCGGAGGACCAACCGACCCAGGAGGTAAAGTGGCC 660
Db 601 CCGGGAATTTGGAGACCCCTGCGGAGGACCAACCGACCCAGGAGGTAAAGTGGCC 660
Qy 661 AGCAACTTATCTGTGCTGCTGATGCTAGTGTCTATGTTTGAATGTTATGCGCTGCG 720
Db 661 AGCAACTTATCTGTGCTGCTGATGCTAGTGTCTATGTTTGAATGTTATGCGCTGCG 720
Qy 721 TCTGTACTAGTTAGTAACTAGTCTGTAATCTGCGGAGCCCGTGGTGGAACTGACGAGTT 780
Db 721 TCTGTACTAGTTAGTAACTAGTCTGTAATCTGCGGAGCCCGTGGTGGAACTGACGAGTT 780
Qy 781 CTGAACACCCCGCGGCAACCCCTGCGGAGACGCTCCAGGGAATTTGGGGCGCGTTTGTGG 840
Db 781 CTGAACACCCCGCGGCAACCCCTGCGGAGACGCTCCAGGGAATTTGGGGCGCGTTTGTGG 840
Qy 841 CCGGACTTGAAGGAGGAGTGCATGGAATCCGAACCCCGTCAAGGATATGTGTTCTGTT 900
Db 841 CCGGACTTGAAGGAGGAGTGCATGGAATCCGAACCCCGTCAAGGATATGTGTTCTGTT 900
Qy 901 AGCAGACGAGAACTTAAACAGTTCCCGCTCTGCTGGAATTTTGTCTTTCGGTTTCGAA 960
Db 901 AGCAGACGAGAACTTAAACAGTTCCCGCTCTGCTGGAATTTTGTCTTTCGGTTTCGAA 960
Qy 961 CCGAAGCCGCGCTCTTGTCTGCTGAGCAACGCTTGGGCTGAGGTCGACTCTAGAGGA 1020
Db 961 CCGAAGCCGCGCTCTTGTCTGCTGAGCAACGCTTGGGCTGAGGTCGACTCTAGAGGA 1020
Qy 1021 TCAATTCGGCACGAGTAAATCGGTGCTGCTTTTAGGACATATGAAGTATGGCAGT 1080
Db 1021 CC-----CTGCACCGGAACATGGAGAACACACATCAGGATTCCT 1059

Db 2460 AAGCAACACAGACAGGTCTGACATGATGGATGGACGAACCACTGAATTCGCAATTGCGAGAT 2519
Qy -ATTGTATTTAAGTGCCTAGCTCGATACAGCAACGCCA--TTTGACCAATTCACCAAT 3357
Db 2520 AATTGTATTTAAGTGCCTAGCTCGATACAGCAACGCCAATTTTGGACCAATTCACCAAT 2579
Qy 3358 GGTGTGCAACCT--CCAAAGCTTCAGCTCCGCAAGCACTCAGGGCGCAAGGGCTGCTAAA 3415
Db 2580 GGTGTGCAACCTTCCAAAGCTTCAGCTCCGCAAGCACTCAGGGCGCAAGGGCTGCTAAA 2639
Qy 3416 GGAAGCGGAACACGTAGAAAGCAGTCCGCAAGAAACGGTGTCTGACCCCGGATGAATGTCA 3475
Db 2640 GGAAGCGGAACACGTAGAAAGCAGTCCGCAAGAAACGGTGTCTGACCCCGGATGAATGTCA 2699
Qy 3476 GCTACTGGGCTATCTGGCAAGGGAACACGCAAGCGCAAGAAAGAGAGACAGGTAGCTTGA 3535
Db 2700 GCTACTGGGCTATCTGGCAAGGGAACACGCAAGCGCAAGAAAGAGAGACAGGTAGCTTGA 2759
Qy 3536 GTGGGCTTACATGGCGATAGCTAGACTGGGCGGTTTATGACACAGCAAGCGGAACCGGAAT 3595
Db 2760 GTGGGCTTACATGGCGATAGCTAGACTGGGCGGTTTATGACACAGCAAGCGGAACCGGAAT 2819
Qy 3596 TGCCAGCTGGGGCGGCTCTGGTAAAGTTGGGAAGCCCTGCAAAAGTAACTGGATGGCTT 3655
Db 2820 TGCCAGCTGGGGCGGCTCTGGTAAAGTTGGGAAGCCCTGCAAAAGTAACTGGATGGCTT 2879
Qy 3656 TCTTGGCGCCAGGATCTGATGGCGGCGGATCAAGATCTGATCAAGACAGGATGAG 3715
Db 2880 TCTTGGCGCCAGGATCTGATGGCGGCGGATCAAGATCTGATCAAGACAGGATGAG 2939
Qy 3716 GATCGTTTCGATGATTTCAACAAAGATGATTTGACGACGAGGTTCTCCGCGGCTTGGGTGG 3775
Db 2940 GATCGTTTCGATGATTTCAACAAAGATGATTTGACGACGAGGTTCTCCGCGGCTTGGGTGG 2999
Qy 3776 AGAGCTATTCCGCTATGACTGGGCACAAAGACAAATCGGCTGCTGATGCCCGCGGT 3835
Db 3000 AGAGCTATTCCGCTATGACTGGGCACAAAGACAAATCGGCTGCTGATGCCCGCGGT 3059
Qy 3836 TCCGGCTGTACGCGGAGGGCGCGGCTCTTTTGTGCAAGACGACTGTCGGTGGCC 3895
Db 3060 TCCGGCTGTACGCGGAGGGCGCGGCTCTTTTGTGCAAGACGAGCTGTCGGTGGCC 3119
Qy 3896 TGAATGAATTCAGACGAGGACGCGGCTATCGTGGCTGCCACACGCGGCTTCTT 3955
Db 3120 TGAATGAATTCAGACGAGGACGCGGCTATCGTGGCTGCCACACGCGGCTTCTT 3179
Qy 3956 GCGCAGCTGTCTCGACGTGTCTAAGCGGGAAGGACTGCTGCTATTTGGCGGAAG 4015
Db 3180 GCGCAGCTGTCTCGACGTGTCTAAGCGGGAAGGACTGCTGCTATTTGGCGGAAG 3239
Qy 4016 TGCCGGGCGAGATCTCTGTATCTACCTTGCCTGCTGCCGAGAAAGTATCCATCATGG 4075
Db 3240 TGCCGGGCGAGATCTCTGTATCTACCTTGCCTGCTGCCGAGAAAGTATCCATCATGG 3299
Qy 4076 GTGATGAATTCGCGGCTGATACGCTTGTATCGGCTACCTGCCCATTCGACCAACCAAG 4135
Db 3300 GTGATGAATTCGCGGCTGATACGCTTGTATCGGCTACCTGCCCATTCGACCAACCAAG 3359
Qy 4136 CGAAACATTCGATCGAGCGACAGTACTCGGATGGAGCGGCTTGTGATCGATCAGGATG 4195
Db 3360 CGAAACATTCGATCGAGCGACAGTACTCGGATGGAGCGGCTTGTGATCGATCAGGATG 3419
Qy 4196 ATCTGGACGAGAGATCAGGGGCTCGCGCCAGCGCAACTGTTGCCAGGCTCAAGGGCG 4255
Db 3420 ATCTGGACGAGAGATCAGGGGCTCGCGCCAGCGCAACTGTTGCCAGGCTCAAGGGCG 3479
Qy 4256 GCATGCCGCGAGGAGATCTGCTGCTGAGCCCATGCGGATGCGCTTGGCCGAATATCA 4315
Db 3480 GCATGCCGCGAGGAGATCTGCTGCTGAGCCCATGCGGATGCGCTTGGCCGAATATCA 3539
Qy 4316 TGGTGGAAAATGGCGGCTTTTCTGGAATTCATCGACTGTGGCGGCTGCGGTGGCGGACC 4375
Db 3540 TGGTGGAAAATGGCGGCTTTTCTGGAATTCATCGACTGTGGCGGCTGCGGTGGCGGACC 3599

Qy 4376 GCTATCAGGACATAGCGTTGGCTACCGGTGATATTGCTGAAGAGCTTGGCGCGCAATGGG 4435
Db 3600 GCTATCAGGACATAGCGTTGGCTACCGGTGATATTGCTGAAGAGCTTGGCGCGCAATGGG 3659
Qy 4436 CTGACCGCTTCTCGTGTCTTTAGCGTATCGCGCTCCCGATTCGACAGCGCATCGCTTCT 4495
Db 3660 CTGACCGCTTCTCGTGTCTTTAGCGTATCGCGCTCCCGATTCGACAGCGCATCGCTTCT 3719
Qy 4496 ATGCGCTTCTGACGAGTCTTCTGACGCGGAGCTCTGGGTTTCGATA 4542
Db 3720 ATGCGCTTCTGACGAGTCTTCTGACGCGGAGCTCTGGGTTTCGATA 3766

RESULT 13
SYNMLPLN1
LOCUS SYNMLPLN1 5464 bp DNA linear SYN 05-JAN-2001
DEFINITION Retroviral vector pLN, complete sequence.
ACCESSION M28245
VERSION M28245.1 GI:208842
KEYWORDS genome; neomycin phosphotransferase; retroviral vector.
SOURCE Retroviral vector pLN
ORGANISM Retroviral vector pLN
REFERENCE 3 (bases 1 to 5464)
AUTHORS Miller, A.D.
TITLE Direct Submission
JOURNAL Submitted (22-SEP-1989) A.D. Miller, Program in Molecular Medicine,
Fred Hutchinson Cancer Research Center, Seattle, WA 98104, USA
REFERENCE 4 (bases 1 to 5464)
AUTHORS Miller, A.D. and Roisman, G.J.
TITLE Improved retroviral vectors for gene transfer and expression
JOURNAL Biotechniques 7 (9), 980-982 (1989)
PUBMED 2631796
FEATURES
Location/Qualifiers
1..5464
/organism="Retroviral vector pLN"
/mol_type="genomic DNA"
/db_xref="taxon:147131"
/notes="derived from Moloney murine leukemia virus"
144..145
/notes="mouse DNA end, Mo-MuSV DNA start"
145..733
/notes="Mo-MuSV 5' long terminal repeat"
803..1612
/notes="extended packaging signal"
1133..1134
/notes="Mo-MuSV DNA end, Mo-MuLV DNA start"
1193..1195
/notes="gag ATG start codon to TAG stop codon"
1616..1617
/notes="Mo-MuLV DNA end, Tn5 DNA start"
1656..2450
/notes="neomycin phosphotransferase"
/codon_start=1
/transl_table=11
/protein_id="AAA72062.1"
/db_xref="GI:208843"
/translation="MIEQDGLHAGSPAAYVRLFGYDWAQQTIGCSDAAVFRLSAQGR
PVLFVKTLSGALNELQDEAARLSWLTATGVPCAAVLDVVTAGRDWLLILGEVPGQDL
LSVHLAPAEKVSIMADAMRRLHTLDPATCFDHOAKHRIERATRMEAGLVDDQDLDE
EHQGLPAELFARLKARMPDGEDLVVTHGDACILPNIENVGSRFSGFTDCGRGLGVADRY
QDIALATRDIAELGGSEWADRFLVLYGIAAPDSQRIAPYRLLEFF"
2467..2468
/notes="Tn5 DNA end, MoMuLV DNA start"
2510..3103
/notes="Mo-MuLV 3' long terminal repeat"
3172..3173
/notes="Mo-MuLV DNA end-plasmid pBR322 DNA start"
misc_recomb
LTR
misc_recomb
ORIGIN
Query Match 28.7%; Score 1487.6; DB 11; Length 5464;
Best Local Similarity 99.7%; Pred. No. 0;

Db	2031		GATCTGATCAAGAGACAGAGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACG	2090
Qy	3752	CAGGTTCTCCGGCCGCTTGGGTGAGAGGCTATTCCGCTATGACTGGGCGACAACAGACAA	3811	
Db	2091	CAGGTTCTCCGGCCGCTTGGGTGAGAGGCTATTCCGCTATGACTGGGCGACAACAGACAA	2150	
Qy	3812	TCGGCTGCTCTGATGCCCGCGTGTTCGGGCTGTGAGCGCAGGGGCGCCCGGTTCTTTTGG	3871	
Db	2151	TCGGCTGCTCTGATGCCCGCGTGTTCGGGCTGTGAGCGCAGGGGCGCCCGGTTCTTTTGG	2210	
Qy	3872	TCAAGACCGACTGTCGCGTGCCCTGAAATGAACTGCAGAGACGAGGACAGCGCGCTATCGT	3931	
Db	2211	TCAAGACCGACTGTCGCGTGCCCTGAAATGAACTGCAGAGACGAGGACAGCGCGCTATCGT	2270	
Qy	3932	GGCTGGGCCACGACGGGGGTTTCCTTGGCCGAGCTGTCTCGAGTTGTTCACATGAAAGCGGAA	3991	
Db	2271	GGCTGGGCCACGACGGGGGTTTCCTTGGCGAGCTGTCTCGAGTTGTTCACATGAAAGCGGAA	2330	
Qy	3992	GGGACTGGCTGCTATTGGGCGAAAGTCCGGGCGAGGATCTCTGTCTCATCTCACCTTGCCTC	4051	
Db	2331	GGGACTGGCTGCTATTGGGCGAAAGTCCGGGCGAGGATCTCTGTCTCATCTCACCTTGCCTC	2390	
Qy	4052	CTGCCGAGAAAGTATCCATCATGTGCTGATGCAATGCGGCGGCTGCATACGTTGATCCGG	4111	
Db	2391	CTGCCGAGAAAGTATCCATCATGTGCTGATGCAATGCGGCGGCTGCATACGTTGATCCGG	2450	
Qy	4112	CTACTCTGCCCATTCGACCCACMACGCGAAACATCGCATCGAGCGGACGCTACTCGGATGG	4171	
Db	2451	CTACTCTGCCCATTCGACCCACMACGCGAAACATCGCATCGAGCGGACGCTACTCGGATGG	2510	
Qy	4172	AAGCCGGTCTTGTTCGATCAGGATGATCTGGAACGAAGAGCATCAGGGGCTCGCGCAGCCG	4231	
Db	2511	AAGCCGGTCTTGTTCGATCAGGATGATCTGGAACGAAGAGCATCAGGGGCTCGCGCAGCCG	2570	
Qy	4232	AATGTTTGGCAGGCTCAAGGCGCGCATGCCCGACGCGAGGATCTCGTCTGACCCCATG	4291	
Db	2571	AATGTTTGGCAGGCTCAAGGCGCGCATGCCCGACGCGAGGATCTCGTCTGACCCCATG	2630	
Qy	4292	GCGATGCTGCTTGCAGGATATCATGTGGAAATGCGCGCTTTCCTGGAATTCATCGACT	4351	
Db	2631	GCGATGCTGCTTGCAGGATATCATGTGGAAATGCGCGCTTTCCTGGAATTCATCGACT	2690	
Qy	4352	GTGGCCGGCTGGGTGTGGCGACCGCTATCAGGACATAGCGTTGGCTACCCGCTGATATTG	4411	
Db	2691	GTGGCCGGCTGGGTGTGGCGACCGCTATCAGGACATAGCGTTGGCTACCCGCTGATATTG	2750	
Qy	4412	CTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCGTCTTACGGTATCGCGCTC	4471	
Db	2751	CTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCGTCTTACGGTATCGCGCTC	2810	
Qy	4472	CCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTCTTCTTGAGCGGGACTCT	4531	
Db	2811	CCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTCTTCTTGAGCGGGACTCT	2870	
Qy	4532	GGGGTTCGATAAAATAAAGATTTTATTAGTCTCCAGAAAAAGGGGGGAATGAAGACC	4591	
Db	2871	GGGGTTCGATAAAATAAAGATTTTATTAGTCTCCAGAAAAAGGGGGGAATGAAGACC	2930	
Qy	4592	CCACTCTAGGTTTGGCAAGCTAGCTTAAAGTAAACCGCAATTTGCAAGGCGATGGAAAAATA	4651	
Db	2931	CCACTCTAGGTTTGGCAAGCTAGCTTAAAGTAAACCGCAATTTGCAAGGCGATGGAAAAATA	2990	
Qy	4652	CATAACTGAGNATAGAGAGTTTCAGATCAAGGTCAGGACAGATGGGAACAGCTGGAATATG	4711	
Db	2991	CATAACTGAGNATAGAGAGTTTCAGATCAAGGTCAGGACAGATGGGAACAGCTGGAATATG	3050	
Qy	4712	GGCCAAAACAGGATATCTGTGTAGCAGTTCCTCCGCCCGCTCAGGGGCCAAGAACAGATG	4771	
Db	3051	GGCCAAAACAGGATATCTGTGTAGCAGTTCCTCCGCCCGCTCAGGGGCCAAGAACAGATG	3110	
Qy	4772	GAAACAGCTGAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTCTCTGCCCGGCTCAG	4831	

Db	3111	GACAGCTGAATATGGGCCAAACAGGANATCTGTGTGAAGCAGTTCTCTGCCCGGCTCAG	3170
Qy	4832	GGCCAGAAACAGATGGTCCCCAGATGCGGTCCAGCCCTCAGCAGTTTCTTAGAGAACCATC	4891
Db	3171	GGCCAGAAACAGATGGTCCCCAGATGCGGTCCAGCCCTCAGCAGTTTCTTAGAGAACCATC	3230
Qy	4892	AGATGTTTCCAGGGTCCCCAAGGACCTGAAATGACCCCTGTGCTTATTTGAACCTAACCA	4951
Db	3231	AGATGTTTCCAGGGTCCCCAAGGACCTGAAATGACCCCTGTGCTTATTTGAACCTAACCA	3290
Qy	4952	ATCAGTTCGGCTCTCCGCTTCTGTTCCGGCGCTTCTGCTCCCGAGCTCAATAAAGAGCC	5011
Db	3291	ATCAGTTCGGCTCTCCGCTTCTGTTCCGGCGCTTCTGCTCCCGAGCTCAATAAAGAGCC	3350
Qy	5012	CACAACCCCTCACTCGGGGGCCAGTCCTCCGATTGACTGAGTCGCCCGGGTACCCCGTGT	5071
Db	3351	CACAACCCCTCACTCGGGGGCCAGTCCTCCGATTGACTGAGTCGCCCGGGTACCCCGTGT	3410
Qy	5072	ATCCAATAAACCTCTTTCAGTTGCACTCCGACTTGCGTGTGCTGCTTCTTGGAGGGT	5131
Db	3411	ATCCAATAAACCTCTTTCAGTTGCACTCCGACTTGCGTGTGCTGCTTCTTGGAGGGT	3470
Qy	5132	CTCTCTGAGTGATTGACTACCCGTCAGCGGGGCTCTTTCATTGG	5177
Db	3471	CTCTCTGAGTGATTGACTACCCGTCAGCGGGGCTCTTTCATTGG	3516
RESULT 15			
SYNMMPLN4			
LOCUS	SYNMMPLN4 5874 bp DNA linear SYN 05-JAN-2001		
DEFINITION	Retroviral vector pLXSN, complete sequence.		
ACCESSION	M28248		
VERSION	M28248.1 GI:208848		
KEYWORDS	genome; neomycin phosphotransferase; retroviral vector.		
SOURCE	Retroviral vector pLXSN		
ORGANISM	Retroviral vector pLXSN		
REFERENCE	3 (bases 1 to 5874)		
AUTHORS	Miller,A.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-SEP-1989) A.D. Miller, Program in Molecular Medicine, Fred Hutchinson Cancer Research Center, Seattle, WA 98104, USA		
REFERENCE	4 (bases 1 to 5874)		
AUTHORS	Miller,A.D. and Roeman,G.J.		
TITLE	Improved retroviral vectors for gene transfer and expression		
JOURNAL	BioTechniques 7 (9), 980-982 (1989)		
PUBMED	2631796		
FEATURES	Location/Qualifiers		
source	1..5874		
misc_recomb	/organism="Retroviral vector pLXSN"		
	/mol_type="genomic DNA"		
LTR	/db_xref="taxon:147134"		
	/notes="derived from Moloney murine leukemia virus"		
misc_signal	174..175		
	/notes="mouse DNA end-Mo-MuSV DNA start"		
misc_recomb	175..763		
	/note="Mo-MuSV 5' long terminal repeat"		
variation	833..1642		
	/notes="extended packaging signal"		
misc_recomb	1163..1164		
	/notes="Mo-MuSV DNA end-Mo-MuLV DNA start"		
misc_signal	1223..1225		
	/notes="gag ATG start codon to TAG stop codon"		
misc_recomb	1634..1655		
	/notes="Mo-MuLV DNA end-simian virus 40 DNA start"		
misc_signal	1655..2020		
	/notes="simian virus 40 early promoter"		
misc_recomb	2020..2021		
	/note="simian virus DNA end-Tn5 DNA start"		
CDS	2066..2860		
	/note="neomycin phosphotransferase"		
misc_recomb	/codon_start=1		
	/transl_table=11		

		/protein_id="AA72065.1"	
		/db_xref="GI:208849"	
		/translation="MTEOGLHAGSPAAWVERLFGYDWAQOTIGCSDAAVFRLSAQGR PVLVFKTDLGALNELQDEARLSWLAATTVPCFAAVLVDVTEAGRWLLILEVPGDL LSHLAPAEKVSIMADAMRRLHTLDPATCPDHOAKHRIERARTRMEAGLVDDQDLDE EHQGLAPAEFLPARLKARMPDGEDLVVTHGDACLPNIMVNGRPSGFIDCGRLGVADRY QDIALATEDIAEELGGEWADRPVLVYGLIAPDSQRIAFYRLLDDEFF"	
misc_recomb		2877*2878	
		/notes="Tn5 DNA end-Mo-MuLV DNA start"	
LTR		2920..3513	
		/notes="Mo-MuLV 3' long terminal repeat"	
misc_recomb		3582*3583	
		/notes="Mo-MuLV DNA end-plasmid pBR322 DNA start"	
ORIGIN			
Query Match		28.7%; Score 1486; DB 11; Length 5874;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 1486; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	3692	GATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACG	3751
DB	2031	GATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACG	2090
QY	3752	CAGGTTCTCGGCGGCTTGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAAGACAA	3811
DB	2091	CAGGTTCTCGGCGGCTTGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAAGACAA	2150
QY	3812	TCGGCTGCTCTGATGCCCGCTGTTCCGGCTGTGACGAGGGCGCCGGTCTCTTTTGG	3871
DB	2151	TCGGCTGCTCTGATGCCCGCTGTTCCGGCTGTGACGAGGGCGCCGGTCTCTTTTGG	2210
QY	3872	TCAAGACCGACCTGTCCGGTCCCTCGAATGAATGCAAGACGAGGCGGCTATCGT	3931
DB	2211	TCAAGACCGACCTGTCCGGTCCCTCGAATGAATGCAAGACGAGGCGGCTATCGT	2270
QY	3932	GGCTGGCCACAGCGGCGTTCCTTGGCGAGCTGTGCTCGACGTTGTCTACTGAAGCGGGAA	3991
DB	2271	GGCTGGCCACAGCGGCGTTCCTTGGCGAGCTGTGCTCGACGTTGTCTACTGAAGCGGGAA	2330
QY	3992	GGGACTGGCTGCTATTGGGCGAAGTGGCGGGCAGGATCTCTGTCTCATCTCACCTTGGCTC	4051
DB	2331	GGGACTGGCTGCTATTGGGCGAAGTGGCGGGCAGGATCTCTGTCTCATCTCACCTTGGCTC	2390
QY	4052	CTGCCGAGAAAGTATCCATCATATGCTGATGCAATGCGGGCGCTGCATACGCTTGATCCGG	4111
DB	2391	CTGCCGAGAAAGTATCCATCATATGCTGATGCAATGCGGGCGCTGCATACGCTTGATCCGG	2450
QY	4112	CTACTCGCCATTTCGACCAACAGCGAAACATCGCATTCGAGCGAGCACGTACTCGGATGG	4171
DB	2451	CTACTCGCCATTTCGACCAACAGCGAAACATCGCATTCGAGCGAGCACGTACTCGGATGG	2510
QY	4172	AAGCCGGTCTTGTGATCAGGATGATCTGGACGAGAGCATCAGGGGCTCGCGCCAGCCG	4231
DB	2511	AAGCCGGTCTTGTGATCAGGATGATCTGGACGAGAGCATCAGGGGCTCGCGCCAGCCG	2570
QY	4232	AACTGTTTCGCGAGGCTCAAGCGCGCATGCCCGACGCGAGGATCTCGCTGACCCCATG	4291
DB	2571	AACTGTTTCGCGAGGCTCAAGCGCGCATGCCCGACGCGAGGATCTCGCTGACCCCATG	2630
QY	4292	GCGATGCTTGTCCGAAATATCATGGTGGAAATGGCCGCTTTTCTGGATTCATCGACT	4351
DB	2631	GCGATGCTTGTCCGAAATATCATGGTGGAAATGGCCGCTTTTCTGGATTCATCGACT	2690
QY	4352	GTGGCGGCTGGGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTG	4411
DB	2691	GTGGCGGCTGGGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTG	2750
QY	4412	CTGAAGAGCTTGGCGGCAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGCCGCTC	4471
DB	2751	CTGAAGAGCTTGGCGGCAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGCCGCTC	2810
QY	4472	CCGATTTCGACGCGCATCGCCTTCTATCGCCTTCTTGACGAGTCTTCTTGAGCGGGACTCT	4531

Db	2811	CCGATTCGACGCGCATCGCCTTCTATCGCCTTCTTGAGAGTCTTCTTGACGGGACTCT	2870
Qy	4532	GGGTTTCGATAAATAAAGATTTTATTAGTCTCCAGAAAAAGGGGGAATGAAAGACC	4591
Db	2871	GGGTTTCGATAAATAAAGATTTTATTAGTCTCCAGAAAAAGGGGGAATGAAAGACC	2930
Qy	4592	CCACCTGTAGTTTGGCAAGCTAGCTTAAGTAAGCCATTTTGAAGCATGGAATAATA	4651
Db	2931	CCACCTGTAGTTTGGCAAGCTAGCTTAAGTAAGCCATTTTGAAGCATGGAATAATA	2990
Qy	4652	CATAACTGAGAATAGAGAAGTTTCAAGTCAAGGTCAGGAACAGATGGAACAGTCAATG	4711
Db	2991	CATAACTGAGAATAGAGAAGTTTCAAGTCAAGGTCAGGAACAGATGGAACAGTCAATG	3050
Qy	4712	GGCCAAACAGGATATCTGTGTAGTTCCTGCCCGGCTCAGGGCCAGAACAGATG	4771
Db	3051	GGCCAAACAGGATATCTGTGTAGTTCCTGCCCGGCTCAGGGCCAGAACAGATG	3110
Qy	4772	GAACAGCTGAATATGGGCCAAAACAGGATATCTGTGTAGTTCCTGCCCGGCTCAG	4831
Db	3111	GAACAGCTGAATATGGGCCAAAACAGGATATCTGTGTAGTTCCTGCCCGGCTCAG	3170
Qy	4832	GGCCAAACAGATGTTCCCGGTCAGGTCGCTCAGCCCTCAGCAGTTCCTAGAGAACCATC	4891
Db	3171	GGCCAAACAGATGTTCCCGGTCAGGTCGCTCAGCCCTCAGCAGTTCCTAGAGAACCATC	3230
Qy	4892	AGATGTTTCCAGGGTGGCCCAAGGACCTGAAATGACCTGTGCCTTATTTGAACTAAACA	4951
Db	3231	AGATGTTTCCAGGGTGGCCCAAGGACCTGAAATGACCTGTGCCTTATTTGAACTAAACA	3290
Qy	4952	ATCAGTTTCGCTTCTGCTTTCGCGCTTCTGCTCCCGAGCTCAATPAAAGAGCC	5011
Db	3291	ATCAGTTTCGCTTCTGCTTTCGCGCTTCTGCTCCCGAGCTCAATPAAAGAGCC	3350
Qy	5012	CACAAACCCCTCACTCGGGGCGCCAGTCTCGGATGACTGATCGCGGGGTACCCGTGT	5071
Db	3351	CACAAACCCCTCACTCGGGGCGCCAGTCTCGGATGACTGATCGCGGGGTACCCGTGT	3410
Qy	5072	ATCCAAATAAACCCCTCTTTGCAAGTTGCATCCGACTTGTGCTCTCGCTTGTGGAGGGT	5131
Db	3411	ATCCAAATAAACCCCTCTTTGCAAGTTGCATCCGACTTGTGCTCTCGCTTGTGGAGGGT	3470
Qy	5132	CTCCTCTGAGTGAATGATACCCGTCAGCGGGGTCTTTTCAATTGG	5177
Db	3471	CTCCTCTGAGTGAATGATACCCGTCAGCGGGGTCTTTTCAATTGG	3516

Search completed: February 27, 2006, 16:11:36
Job time : 25052 secs

THIS PAGE BLANK (USPTO)